Article

PERCEPTION

Perception
2015, Vol. 44(8–9) 1064–1076

DOI: 10.1177/0301006615596872
pec.sagepub.com
SAGE

Scanpaths of Complex Image Viewing: Insights From Experimental and Modeling Studies

Anatoly Samarin, Tatiana Koltunova, Vladislav Osinov, Dmitry Shaposhnikov and Lubov Podladchikova
A.B. Kogan Research Institute for Neurocybernetics, at Southern Federal University, Rostov-on-Don, Russia

Abstract
From the first works of Buswell, Yarbus, and Noton and Stark, the scan path for viewing complex images has been considered as a possible key to objective estimation of cognitive processes and their dynamics. However, evidences both pro and con were revealed in the modern research. In this article, the results supporting the Yarbus-Stark concept are presented. In psychophysical tests, two types of images (three paintings from Yarbus' works and four textures) were used with two instructions, namely, "free viewing" and "search for modified image regions." The focus of the analysis of experimental results and modeling has been given to local elements of the scan path. It was shown that each parameter used (square of viewing area, *S*; distance between center of mass of viewing area and image center, *R*; parameter X_i, based on duration of the current fixation and angle between preceding and following saccades), reflects the specificity of both visual task and image properties. Additionally, the return gaze fixations which have a set of specific properties and mainly address to the areas of interest on image were revealed. Evidently these facts can be formalized in an advanced mathematical model as additional instrument to study the mechanisms of complex image viewing.

Keywords
visual attention, scan path, fixation duration, areas of interest, return fixations, psychophysical experiment, mathematical modeling

Introduction

From the first works of Buswell (1935), Yarbus (1967), and Noton and Stark (1971), the scan path of complex image viewing has been considered as a possible key to objective estimation of cognitive processes and their dynamics. However, evidence both for (Borji & Itti, 2014;

Corresponding author:
Lubov Podladchikova, A.B. Kogan Research Institute for Neurocybernetics, Academia of Biology and Biotechnology, Southern Federal University, Rostov-on-Don, Russia.
Email: lnp@nisms.krinc.ru

DeAngelus & Pelz, 2009; Harding & Bloj, 2010; Humphrey & Underwood, 2010; Oyekoya & Stentiford, 2007; Priviterra & Stark, 2005; Tatler, Wade, Kwan, Findlay, & Velichkovsky, 2009) and against (Greene, Liu, & Wolfe, 2012) has been revealed in modern research. Several works, substantially confirming the Yarbus-Stark concept, found experimental conditions for which the specificity of scan path is less obvious (DeAngelus & Pelz, 2009). These authors consider possible reasons to receive various results, namely, peculiarities of the instructions before the tests, test duration, subject report after the tests and, most of all, mathematical methods used for quantitative analysis of the scan path. In particular, the analysis of initial experimental dataset from Greene et al. (2012) by different nonlinear methods, such as MultiFixation Pattern Analysis, Fisher Kernel Learning Algorithm, K-Nearest-Neighbor method (Borji & Itti, 2014; Kanan, Ray, Bseison, Hsiao, & Cottrell, 2014), shown a meaningful connection between eye movement patterns and visual tasks as determined by the instructions given to participants. It had been shown that viewing scan paths contain information about image encoding, recognition, and classification (Hayes, Petrov, & Sederberg, 2011; Humphrey & Underwood, 2010; Oyekoya & Stentiford, 2007; Weger, Abrams, Law, & Pratt, 2008; Wang & Theeuwes, 2012; West, Haake, Rozanski, & Karn, 2006).

Contrary to detailed investigations of viewing scan path as a whole, there are few works that have been devoted to the study of the local elements of the eye movement patterns (Pannasch, Schulz, & Velichkovsky, 2011; Pastukhov & Braun, 2010). The search for eye movement parameters which allow us to estimate the visual task to be solved during the current stage of dynamical image processing, and evaluation of the contribution of dominating components of visual attention is to be unsolved objectives up to now in both experimental and modeling studies (Carrasco, 2011; Lupianez, Klein, & Bartolomeo, 2006; Navalpakkam & Itti, 2005; Walther & Koch, 2007; Wang & Theeuwes, 2012; Wolfe, Birnkrant, Kunar, & Horowitz, 2005; Zelinsky, 2005).

This article describes the results supporting the idea about viewing scan path as a marker of visual task. The focus of the analysis of experimental results and modeling has been given to local elements of the scan path, in particular to the search for ''return fixations'' (RFs) on recently viewed image regions.

Methods

Participants

Two series of psychophysical experiments were carried out: (a) free viewing of complex images and search for their modified regions; each of 12 volunteers participated in the both types of this experiment; the age of the subjects ranged from 21 to 45; (b) free viewing of textures; three volunteers participated in this experiment; the age of the subjects ranged from 24 to 35. All subjects were naive to the purpose of the experiments, had normal or corrected-to-normal vision, and did not report any neurological or psychological disease. The Bioethics Committee of Southern Federal University approved experimental protocol. All volunteers signed the agreement to participate in experiment.

Stimuli

In Experiment 1, three pictures (Figure 1(a)) from the work of Yarbus (1967) were used: I. E. Repin, ''Unexpected Return'' (Im 1); I. I. Shishkin, ''Countess Mordvinov's Forest'' (Im 2); I. I. Levitan, ''Birch grove'' (Im 3). In Experiment 2, four textures were used

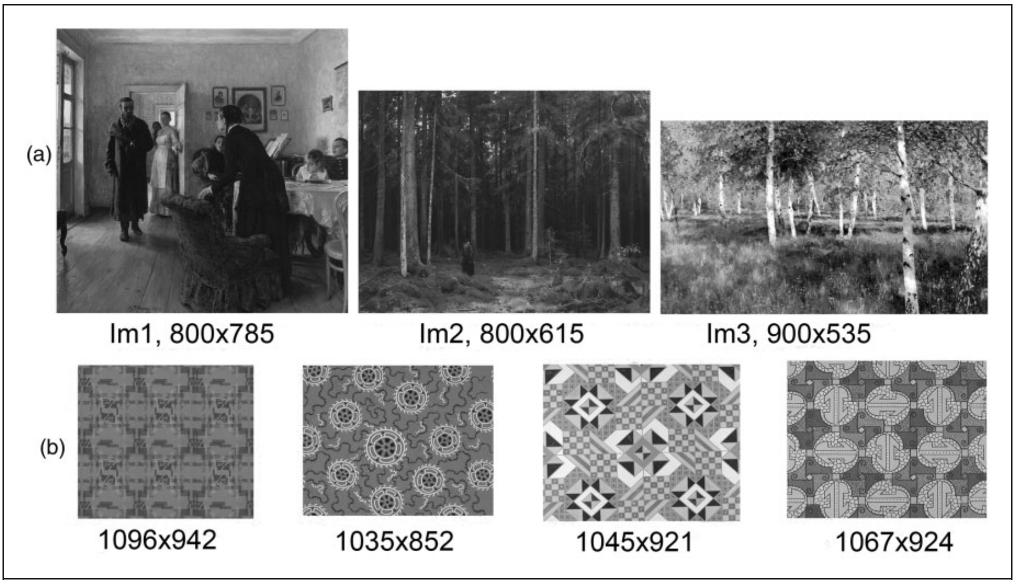

Figure 1. Test images used in two series of psychophysical experiments: (a) Three pictures from work of Yarbus (1967); (b) Four texture images. Image size in pixels is indicated under each picture.

(Figure 1(b)); they had regular spatial structure of elements and different primary features (Osinov, Shaposhnikov, Koltunova, & Podladchikova, 2012).

Procedure

In Experiment 1, for each subject, images from Im 1 to Im 3 were presented with instructions: "free viewing" and "search for modified regions in the same images." Unmodified initial images were used in "free viewing" tests. The diameter of square-modified image regions, blurred by Gauss transformation (Podladchikova, Shaposhnikov, Tikidgji-Hamburyan, et al., 2009), was equal to 2°. Each trial was terminated by subject stated that "image is viewed, I can describe it" or "all modified image regions are found"; duration of viewing of each picture was varied, in the most tests from 1 to 3 minutes. In Experiment 2, each texture was presented during 30 seconds, duration of these tests was determined by experimenter to receive a number of gaze fixations (120 ± 11) considered to be enough to build the detailed viewing trajectory.

Eye Movements Recording and Data Analysis

Eye movements were recorded by SMI iView X Hi-Speed eye-tracker with frame rate of 1250 Hz. The distance between subject's eyes and computer screen was 80 cm in Experiment 1 and 50 cm in Experiment 2. Frame rate of stimulus monitor (NEC MultiSync LCD 1990Sxi) is equal to 60 Hz, monitor resolution is equal to 1280×1024 pixels. Saccades and gaze fixations were detected automatically online by the iView X and our developed program implemented by EventIDE software (http://okazolab.com/) The threshold of eye movement velocity was chosen as 40°/s to detect the saccades of small amplitude. All gaze fixations with duration less than 80 ms as well as fixations, which had variance higher than 1° (6% from fixations detected in all trials), were excluded from further analysis. Statistical analysis was performed by BeGaze software and R: A Language and Environment for Statistical Computing (http://www.R-project.org). Differences between particular samples of data were evaluated by the Wilcoxon signed-rank test and the Student *t* test.

Results and Discussion

Spatio-Temporal Properties of Return Fixations

The search for fixations that return the gaze to recently viewed image regions has been performed because they can reflect the high-level processes of visual information processing. A hypothesis about gaze RFs emerged from quantitative analysis of experimental data (Podladchikova, Shaposhnikov, Tikidgji-Hamburyan, et al., 2009), in particular the object-return trajectories identified (Figure 2). Such trajectories of eye movements were recorded while subjects performed particular visual tasks, determined by instruction, or during some stages of free image viewing.

Search for the RFs was performed using the results of Experiments 1 and 2. A fixation was considered to be a RF (point "3" in Figure 3(a)) if it was located in the foveal vicinity ($r = 2°$) of a preceding fixation (point "1" in Figure 3(a)).

Probability (*p*) and spatial location of RFs as well as fixation duration in the vicinity of RFs (separately for points "1," "2," and "3" on the scheme, Figure 3(a)), were determined. The *p* of RFs was calculated as their proportion of the total number of fixations (both regular and the RFs) in each test, or into specific image areas. The average data presented in Table 1 indicate that the *p* of RFs was low for all test images. In the particular tests, it varied from 0.03 up to 0.19.

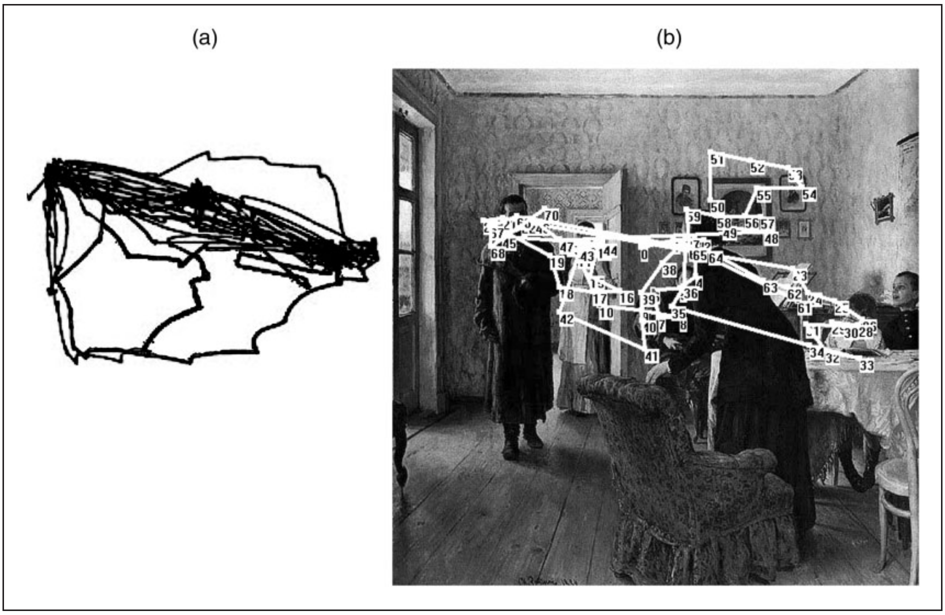

Figure 2. Examples of object-returned scan paths. (a) Yarbus' results (1965), a part of Figure 109; (b) our experimental results.

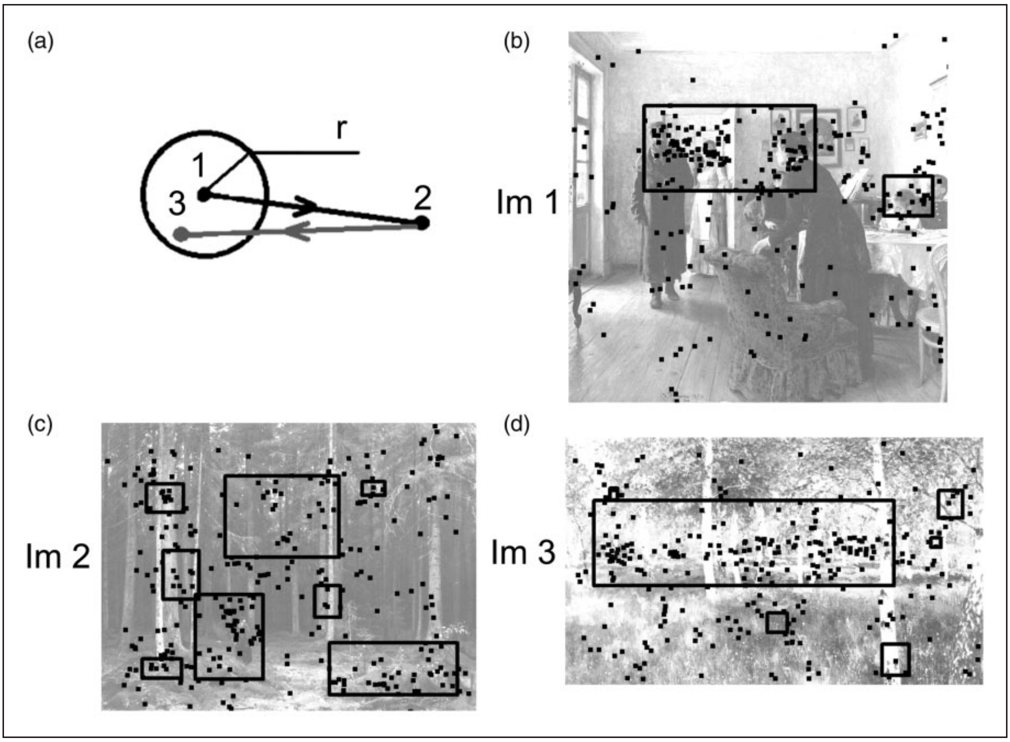

Figure 3. The scheme for determination of the RFs (a) cumulative maps of RF (black small squares) distribution for all subjects; (b–d) the areas of interest identified by analysis of spatial distribution of all fixations (both regular and RFs) are marked by black rectangles.

Table 1. Probability of Gaze Return Fixations During Viewing of Test Images.

Images (Figure 1)	Im 1	Im 2	Im 3	Textures
Total number of fixations, *n*	2,266	2,178	2,380	776
Probability of return fixations, *p*	0.10 ± 0.01	0.12 ± 0.01	0.11 ± 0.01	0.12 ± 0.02

In spite of low probability of RFs, they strongly adress to the Areas of Interest (AOIs, or the most informative regions according to term of Yarbus, 1967). The AOIs were identified by quantitative analysis of distribution of all fixations (both regular and the RFs ones) using modified method of the nearest neighbor (Podladchikova, Shaposhnikov, Tikidgji-Hamburyan, et al., 2009). In Figure 3(b–d), it can be seen that most of the RFs (in average 75%) are located inside AOIs. Besides, the density of the RFs is significantly higher ($p < .05$, the Student t test) inside the AOIs than in other image regions (Table 2).

Duration of fixations in their consequences (points "1," "2," and "3"in Figure 3(a)), including the RFs, was also analyzed. It was revealed that duration of RF is longer compared with the first fixations and those that is not the part of such consequences (Table 3). These differences are significant only for Im 1 ($p < .01$, Wilcoxon signed-rank test).

Table 2. The Averaged Density of the RFs Inside and Outside AOIs.

	Density of the RFs ($n/1°$)	
Images	Inside AOIs	Outside AOIs
Im 1	0.29 ± 0.09 ($n = 130$)	0.04 ± 0.01 ($n = 134$)
Im 2	0.23 ± 0.04 ($n = 147$)	0.08 ± 0.02 ($n = 163$)
Im 3	0.34 ± 0.06 ($n = 155$)	0.08 ± 0.02 ($n = 136$)

Note. RFs = return fixations; AOIs = areas of interests.

Table 3. Averaged Duration of Fixations in the Consequences of Points "1," "2," "3" (Figure 3(a)).

	Location of fixation in their consequences			
Images	First fixation	Second fixation	Third (return) fixation	Fixations outside "1 to 3" consequences
Im 1	519 ± 24 ms ($n = 265$)	524 ± 22 ms ($n = 265$)	568 ± 26 ms ($n = 265$)	428 ± 15 ms ($n = 509$)
Textures	292 ± 20 ms ($n = 63$)	360 ± 26 ms ($n = 63$)	358 ± 35 ms ($n = 63$)	333 ± 26 ms ($n = 63$)

Table 4. Temporal Dynamics of the p and Duration of RFs During Experiment 1.

Period number	First	Second	Middles	Penultimate	Last
Total number of fixations	405	406	1,730	310	404
Probability of return fixations, p	.08	.10	.13	.11	.07
Duration of return fixations, ms	321 ± 25	561 ± 82	433 ± 22	537 ± 110	439 ± 69

Note. RFs = return fixations.

The temporal dynamics of RF duration during a test was analyzed in the dataset obtained in Experiment 1 by methods developed earlier (Podladchikova, Shaposhnikov, Koltunova, Dyachenko, & Gusakova, 2009). Namely, the consequent periods (from initial test stages up to final ones) were determined in the data of each experiment (260 periods in all tests). Each period includes 30 gaze fixations (both the regular and the RFs). Then the average data of all experiments were calculated, and the p and duration of RFs in each period determined. It was revealed that these RF parameters varied from one period to the other (Table 4). In particular, duration of RFs in the first, middle, and last periods is significantly less ($p < .01$, Wilcoxon signed-rank test) than in neighbor periods.

Thus RFs have relatively low probability (up to .19), but peculiarities of such local scan path elements differentiate them from regular fixations, namely, (a) RFs mainly address to the AOI; (b) duration of RFs was more than regular fixations during viewing of both complex pictures and simple images (textures); (c) RF duration varied during image viewing from initial stages up to final ones.

The obtained results can be compared with known evidences achieved according to *Inhibition of Return* (Lupianez et al., 2006; Posner, Rafal, Choate, & Vaughan, 1985; Wang, Satel, & Klein, 2012; Weger et al., 2008) and *Facilitation of Return* (Dodd, Van der

Stigchel, & Hollingworth, 2009; Hooge, Over, van Wezel, Maarten, & Frens, 2005; Luke, Schmidt, & Henderson, 2013) concepts. In particular, the difference between fixation duration of regular and the RFs is in agreement with known data (Bays & Husain, 2012; Hooge et al., 2005; Lupianez et al., 2006). Such difference in fixation duration may reflect a short-term memory processes to be activated for detailed analysis of image fragments located in the AOIs. Other methods to identify the RFs (saccades) at viewing of complex images were used in the known studies. However, *p* of RFs, revealed in different works, was also low. In particular, *p* of RFs in work of Luke et al. (2013) varied around average value (.08 from data presented in Figure 3 from Luke et al., 2013). Similar data were indicated in work of Hooge et al. (2005), *p* varied from .02 to .12 in tests of viewing and search. It was proposed that *Inhibition of Return* mechanisms dominated during search and viewing of simple images (Hooge et al., 2005; Lupianez et al., 2006; Wang & Theeuwes, 2012; Weger et al., 2008). On the contrary, *Facilitation of Return* was often revealed during viewing complex images and scenes (Dodd et al., 2009; Luke et al., 2013).

The results about temporal dynamics of the RF duration are in agreement with known data about temporal dynamics of duration of the regular fixations at the initial test stages (Unema, Pannasch, Joos, & Velichkovsky, 2005) and at the last test stage (Podladchikova, Shaposhnikov, Koltunova, et al., 2009). These evidences indicate that the preattentive mechanisms may dominate in generation of the RFs at the first and the last (before making the decision by subject to finish the test) stages of image viewing. It is possible that the high-level mechanisms may determine RF generation at other viewing periods.

Evidently, the problem of RFs must be investigated in details by experimental and modeling methods because it may give an important key to quantitative estimation of visual attention dynamics. In particular, in most of biologically motivated models of image viewing (Navalpakkam & Itti, 2005; Rybak, Gusakova, Golovan, Podladchikova, & Shevtsova, 2005; Walther & Koch, 2007), it is introduced an empirical coefficient for inhibition of return to prevent cycles in the model scan paths and does not take into account the possibility of facilitation of return. Formalization of relationship between facilitation and inhibition of return allows us to develop the realistic model of viewing scan paths as additional instrument to study the mechanisms of visual attention dynamics. To reach this goal, a correct quantitative comparison of experimental and modeling results is necessary.

Quantitative Estimation of Local Elements of Viewing Scan Path

Several methods (Osinov et al., 2012) were used to analyze the local elements of viewing scan path to be recorded in psychophysical tests. Namely, the quantitative characteristics of image viewing trajectory are as follows.

(1) *S* is the square of viewing area. In each test, the viewing area was determined as image part inside contour figure (see Figure 7) to be formed by external fixation points. *S* of viewing area calculated as a percentage from the entire image area.
(2) *R* is distance between center of mass of viewing area and image center.
(3) The parameter X_i is based on the duration of the current fixation and the angle between saccades preceding and following it (Figure 4(a)). The last characteristic was received as follows. At first, the description of each fixation point of viewing trajectory was obtained and estimated its belonging to one of 49 classes of X_i values. This number of classes was chosen according to preliminary evaluation of variation range of both fixation duration (t_i) and angle between two adjacent saccades (φ_i). In particular, the class "1" of X_i

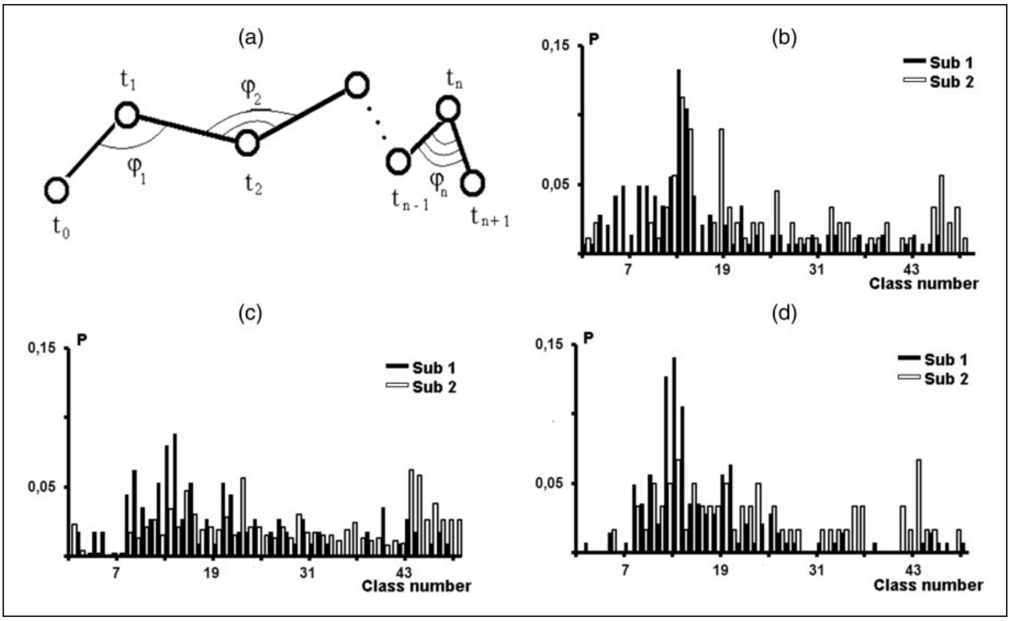

Figure 4. Scheme (a) for estimation of X_i parameter to evaluate the spatio-temporal structure of local scan path elements; fixation points are marked by circles, saccades are marked by lines, t_i—fixation duration (ms), φ_i—angle between two adjacent saccades. Examples of distributions of X_i parameter ((b)–(d)): at viewing texture images (b, $r = 0.52$); free viewing of complex images (c, $r = 0.17$); and search for their modified regions (d, $r = 0.53$).

included the following range of parameters: t_i up to 100 ms, φ_i up to 15°; the class "49" had maximal values of the parameters: t_i more than 1100 ms; φ_i 165 ÷ 180° (see details in Osinov et al., 2012). Finally, a description of each scan path as a whole was evaluated by frequency distribution of X_i values among 49 classes (e.g., Figure 4(b–d)).

The parameters described above were calculated using the results of Experiments 1 and 2. It was revealed that (a) S (percentage of viewing area square from the entire image square) for different volunteers and tasks varied from 3% to 75% for complex images and from 31% to 56% for texture images; (b) R varied from 0.42° to 8.42° for complex images and from 0.59° to 1.75° during viewing of textures. In other words, coordinates of center mass for viewing areas were located near the image center and had low variability during viewing of textures. Both facts as well as small fixation duration (see Table 3) indicate dominance of the low-level mechanisms while processing these simple images. Besides, it was shown that the values of S and R for the same subjects ($n = 12$) depended on the test type while presentation of complex images. In average, S is equal to $43 \pm 2\%$ (coefficient of variation = 28%, between different subjects) in search tests and $34 \pm 3\%$ (coefficient of variation = 42%) in free viewing tests, R is equal to $2.74 \pm 0.28°$ in search tests and $4.20 \pm 0.31°$ in free viewing tests.

Quantitative comparison of viewing scan paths for different volunteers was performed also by Pearson's coefficient correlation r between distributions of parameter X_i. It was revealed that r varied in relatively narrow range at viewing textures (from 0.52 to 0.82). At viewing

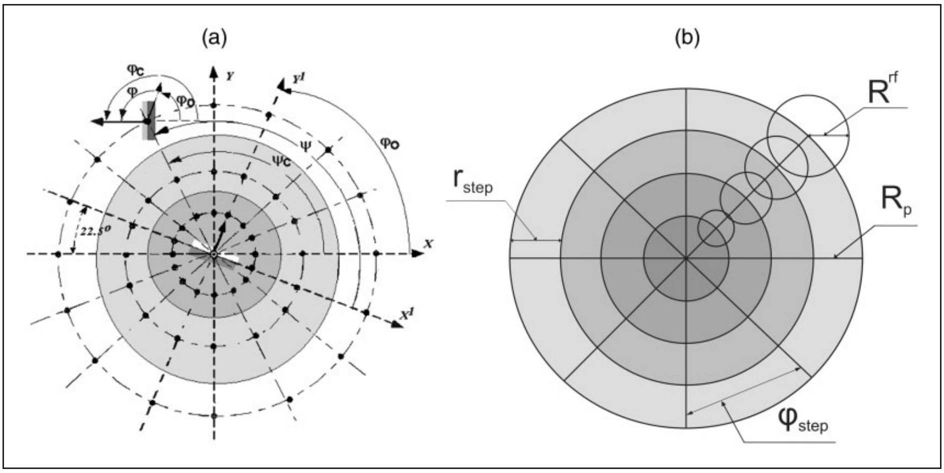

Figure 5. The scheme of a space-variant ("foveal") model input window. (a) the primary version, Figure 2 from Rybak et al. (2005); (b) the current version, R_p—size of input window as a whole; R_{rf}—size of context area of each node of input window; r_{step} and φ_{step}—parameters for concentric and radial structure of input window.

complex images, r was changed in wider range (from 0.15 to 0.81). Besides, the difference between distributions of X_i parameter for the same pairs of subjects was more expressed in free viewing tests as compare with search tests (averaged coefficients of correlation are equal to 0.45 ± 0.07 and 0.54 ± 0.05, correspondingly). The example of distributions of X_i parameter presents in Figure 4 (compare parts "c" and "d"). The results obtained indicate possibility of quantitative comparison of scan paths (both trajectory as a whole and their local elements) while solution of different visual tasks and viewing of different images by means of the used parameters (S, R, and X_i).

Dynamics of Model Scan Path While Changing Input Window Structure

Our mathematical model developed earlier (Podladchikova, Shaposhnikov, Tikidgji-Hamburyan, 2009; Rybak et al., 2005) as additional instrument to study the mechanisms of complex image viewing includes space-variant ("foveal") input window (Figure 5), imitating the visual acuity changing from the center to the periphery of the human visual field, and gaze attraction function. The last function determines chose of image fragments for consequent fixations of input window and formation of model scan path as a whole. The first model version similar to the most known models (Navalpakkam & Itti, 2005; Priviterra & Stark, 2005; Walther & Koch, 2007; Zelinsky, 2005) was mainly based on primary features or saliency map.

Similar models can reproduce basic features (without details) of experimental viewing scan path (Figure 6(c)) and analyze a contribution of the low-level mechanisms while visual information processing. However, search for parameters to include the high-level mechanisms into model gaze attraction function is unsolved task up to now (Walther & Koch, 2007). One of the approaches to this goal can consist in detailed model-based investigation of low-level parameters contribution in scan path formation.

For this goal, the previously published basic model (Podladchikova, Shaposhnikov, Tikidgji-Hamburyan, et al., 2009) was modified to establish the dependence of viewing

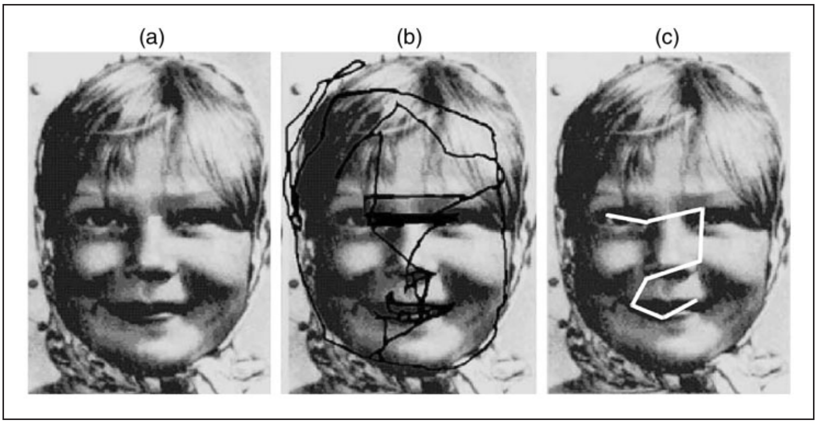

Figure 6. Qualitative comparison of experimental and model viewing scan paths. (a) test image (the photo "Volghanochka," S. Fridland); (b) eye movement trajectory recorded in Yarbus' experiment (1967, modified Figure 115); (c) simulation of image viewing scan path by our feature-based model.

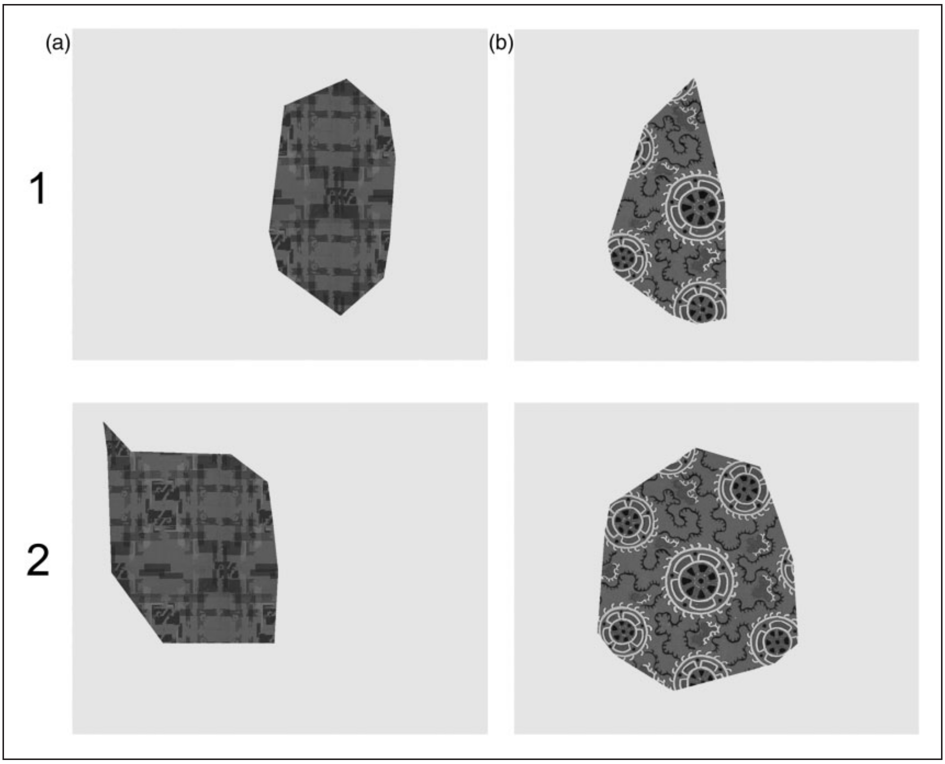

Figure 7. Samples of dynamics of location (a) and square of image scanning area (b) in dependence on R_p, (1) 4°, (2) 8°.

topology on the parameters of input window, that is, R_p, R_{rf}, r_{step}, and φ_{step} (Figure 5(b)). The same texture images as in psychophysical study (Figure 1(b)) and the same parameters (i.e., R and S) were used to analyze a topology of artificial scan paths generated by the model during computer experiments. To correct comparison of the results of different computer experiments, the number of fixation points of input window was equal to 200, and image center was chosen as initial point of "viewing" in all simulation tests. It revealed that decrease of the value of φ_{step} from 72° to 18° resulted in decrease of S in 1.5 times. Values of R_p were chosen to be equivalent to 8°, 12°, and 16° of the human vision field. It was shown that S of viewing area was increased (in average, for all images in 1.5 times) at increasing input window size (Figure 7).

Some results obtained during studying dynamics of model scan path in computer experiments may be compared with the known data of psychophysical research. In particular, dynamics of square for scanning area (parameter S) during "viewing" textures revealed at changing the size of the model input window is in agreement with data about the dynamics of scan path topology of complex image viewing at varying human vision field size by means of special tool (Figure 2, in Rozhkova & Yarbus, 1977). This agreement of experimental and modeling results allows us to suppose an important contribution of the mechanisms of primary sensory level in phenomena described by Rozhkova & Yarbus (1977).

Conclusion

The results of our experimental and modeling studies can be summarized as follows. It was shown that each used parameter for quantitative evaluation of viewing scan path (distance between center mass of viewing area and image center; parameter based on fixation duration and angle between preceding and following saccades; viewing region square) reflects the specificity of visual task, image properties and, evidently, individual differences. In particular, variability of all topology characteristics at viewing complex images between different subjects was more than at viewing relatively simple texture images. Similar differences were revealed between free viewing tests and search for modified regions of the same complex images. These results allow us to suppose that individual peculiarities of viewing strategy, evidently, to be determined by the high-level mechanisms of visual perception are decreased at viewing simple images and solution of search task.

Several obtained results are important to develop the advanced mathematical model of image viewing. They are as follows: (a) trend to decrease the RF duration in the last viewing period in tests terminated by subject; (b) RFs mainly addressed to image AOI; (iii) pronounced dynamics of chosen quantitative parameters of image viewing topology. It is supposed that these phenomena allow us to evaluate a contribution of mechanisms of overt and covert visual attention and receive quantitative criteria to estimate a visual task under solution during the current stage of image viewing.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The work is supported by the SFedU grant № 213.01-2014/001.

References

Bays, P., & Husain, M. (2012). Active inhibition and memory promote exploration and search of natural scenes. *Journal of Vision*, *12*, 1–18.

Borji, A., & Itti, L. (2014). Defending Yarbus: Eye movements reveal observers' task. *Journal of vision*, *14*, 1–29.

Buswell, G. T. (1935). *How people look at pictures*. Chicago: University of Chicago Press.

Carrasco, M. V. (2011). Visual attention: The past 25 years. *Vision Research*, *51*, 1484–1525.

DeAngelus, M., & Pelz, J. B. (2009). Top-down control of eye movements: Yarbus revisited. *Visual Cognition*, *17*(6–7), 790–811.

Dodd, M. D., Van der Stigchel, S., & Hollingworth, A. (2009). Novelty is not always the best policy. Inhibition of return and facilitation of return as a function of visual task. *Psychological Science*, *20*, 333–339.

Greene, M. R., Liu, T., & Wolfe, J. M. (2012). Reconsidering Yarbus: A failure to predict observers' task from eye movement patterns. *Vision Research*, *62*, 1–8.

Harding, G., & Bloj, M. (2010). Real and predicted influence of image manipulations on eye movements during scene recognition. *Journal of Vision*, *10*, 1–17.

Hayes, T. R., Petrov, A. A., & Sederberg, P. B. (2011). A novel method for analyzing sequential eye movements reveals strategic influence on Raven's Advanced Progressive Matrices. *Journal of Vision*, *11*, 1–11.

Hooge, I. Th. C., Over, E. A. B., van Wezel, R. J. A., Maarten, A., & Frens, M. A. (2005). Inhibition of return is not a foraging facilitator in saccadic search and free viewing. *Vision Research*, *45*, 1901–1908.

Humphrey, K., & Underwood, G. (2010). The potency of people in pictures: Evidence from sequences of eye fixations. *Journal of Vision*, *10*, 1–10.

Kanan, C., Ray, N. A., Bseison, D. N. F., Hsiao, J. H., & Cottrell, G. W. (2014). Predicting an observer's task using multi-fixation pattern analysis. In *Proceedings of the ACM symposium on eye tracking research and applications* (pp. 287–290). New York, NY, USA: ACM.

Luke, S. G., Schmidt, J., & Henderson, J. M. (2013). Temporal oculomotor inhibition of return and spatial facilitation of return in a visual encoding task. *Frontiers in Psychology*, *4*, 1–8.

Lupianez, J., Klein, R. M., & Bartolomeo, P. (2006). Inhibition of return: Twenty years after. *Cognitive Neuropsychology*, *23*, 1003–1014.

Navalpakkam, V., & Itti, L. (2005). Modeling the influence of task on attention. *Vision Research*, *45*, 205–231.

Noton, D., & Stark, L. (1971). Scanpaths in eye movements during pattern perception. *Science*. *New Series*, *171*, 308–311.

Osinov, V. A., Shaposhnikov, D. G., Koltunova, T. I., & Podladchikova, L. N. (2012). Dynamics of spatio-temporal features of image viewing: Model and experiment. *Electronic peer-review journal Neuroinformatics*, *6*, 1–11. (in Russian).

Oyekoya, O., & Stentiford, F. (2007). Perceptual image retrieval using eye movements. *International Journal of Computer Mathematics*, *84*, 1379–1391.

Pannasch, S., Schulz, J., & Velichkovsky, B. M. (2011). On the control of visual fixation durations in free viewing of complex images. *Attention, Perception, & Psychophysics*, *73*, 1120–1132.

Pastukhov, A., & Braun, J. (2010). Rare but precious: Microsaccades are highly informative about attentional allocation. *Vision Research*, *50*, 1173–1184.

Podladchikova, L. N., Shaposhnikov, D. G., Koltunova, T. I., Dyachenko, A. V., & Gusakova, V. I. (2009). Temporal dynamics of fixation duration, saccade amplitude, and viewing trajectory. *Journal of Integrative Neuroscience*, *8*, 487–501.

Podladchikova, L. N., Shaposhnikov, D. G., Tikidgji-Hamburyan, A. V., Koltunova, T. I., Tikidgji-Hamburyan, R. A., Gusakova, V. I., & Golovan, A. V. (2009). Model-based approach to study the mechanisms of complex image viewing. *Journal of Optical Memory and Neural Networks (Information Optics)*, *18*, 114–121.

Posner, M. I., Rafal, R. D., Choate, L. S., & Vaughan, J. (1985). Inhibition of return: Neural basis and function. *Cognitive neuropsychology*, *2*, 211–228.

Priviterra, C. M., & Stark, L. (2005). Scanpath theory, attention and image processing algorithms for prediction of human eye fixations. In L. Itti, G. Rees, & J. K. Tsotsos (Eds.), *Neurobiology of attention* (pp. 296–299). New York, NY: Elsevier, Academic.

Rozhkova, G. I., & Yarbus, A. L. (1977). Features of object perception by human in limited areas of the central retina. *Human Physiology*, *3*, 1119–1127. (in Russian).

Rybak, I., Gusakova, V., Golovan, A., Podladchikova, L., & Shevtsova, N. (2005). Attention-guided recognition based on «What»; and «Where» representations: A behavioral model. In L. Itti, G. Rees, & J. K. Tsotsos (Eds.), *Neurobiology of Attention* (pp. 663–670). New York, NY: Elsevier, Academic.

Tatler, B. W., Wade, N. J., Kwan, H., Findlay, J. M., & Velichkovsky, B. M. (2009). Yarbus, eye movements, and vision. *i-Perception*, *1*, 7–27.

Unema, P. J., Pannasch, S., Joos, M., & Velichkovsky, B. M. (2005). Time course of information processing during scene perception: The relationship between saccade amplitude and fixation duration. *Visual Cognition*, *12*, 473–494.

Walther, D. B., & Koch, C. (2007). Attention in hierarchical models of object recognition. *Progress in Brain Research*, *165*, 57–78.

Wang, Z., Satel, J., & Klein, R. M. (2012). Sensory and motor mechanisms of oculomotor inhibition of return. *Experimental Brain Research*, *218*, 441–453.

Wang, Z., & Theeuwes, J. (2012). Dissociable spatial and temporal effects of Inhibition of Return. *PloS One*, *7*, e44290.

Weger, U. W., Abrams, R. A., Law, M. B., & Pratt, J. (2008). Attending to objects: Endogenous cues can produce inhibition of return. *Visual Cognition*, *16*, 659–674.

West, J. M., Haake, A. R., Rozanski, E. P., & Karn, K. S. (2006). eyePatterns: Software for identifying patterns and similarities across fixation sequences. In K.-J. Raiha, A. T. Duchowski (Eds.), *Proceedings of the 2006 symposium on eye tracking research & applications* (pp. 149–154). San Diego, USA.

Wolfe, J. M., Birnkrant, R. S., Kunar, M. A., & Horowitz, T. S. (2005). Visual search for transparency and opacity: Attentional guidance by cue combination? *Journal of Vision*, *5*, 257–274.

Yarbus, A. L. (1967). In L. A. Rigss (Ed.), *Eye movements and vision*. New York, NY: Plenum Press.

Zelinsky, G. J. (2005). Specifying the components of attention in a visual search task. In L. Itti, G. Rees, & J. K. Tsotsos (Eds.), *Neurobiology of attention* (pp. 395–400). New York, NY: Elsevier, Academic.