SMOOTH PURSUIT IN HUMAN VISION AND POSSIBILITIES OF ITS SIMULATION

Valentina P. Bozhkova
Nadezhda S. Surovicheva
Dmitry P.Nikolaev
Dmitry G. Lebedev
Institute for Information Transmission Problems, RAS
Bol'shoj Karetnyj lane, 19, 127994, Moscow, Russia
E-mail: bgk@iitp.ru; ncc@iitp.ru; dimonstr@iitp.ru; dgl@iitp.ru

KEYWORDS
Human vision, eye movement, smooth pursuit

ABSTRACT

Possible advantages of bionic models based on the algorithms of human vision are discussed. Experimentally it was demonstrated that when observing moving objects the human vision system estimates the dynamic characteristics of the object and can program smooth pursuit on the basis of those. The programmed pursuit considerably widens the capabilities of the foveal vision system.

INTRODUCTION

Due to narrow foveal vision, the human vision system has a unique feature: it ensures a high visual acuity within a wide viewing angle. At the same time, the visual acuity of a static eye decreases 3 times when the object image is shifted with respect to the best resolution position by only 2 arc degrees. Should a small object move across the field of view with a speed of a few degrees per second (which is common in practice), the object will promptly leave the visual acuity zone unless the visual axis turns. For perceiving a moving object with the sharp-eyed foveal vision system, people use special oculomotor tracing named smooth pursuit (SP). The purpose of the SP is to stabilize the image of the moving object formed on the retina in order to perceive it in more detail. The SP is a remarkable example of a “computerized” vision control system in biology.

SMOOTH PURSUIT PHENOMENON

To make the SP system switch on, the velocity of the moving object should range from 1 to 50 deg/sec (Yarbus 1965). Shortly after the onset of the object motion, the eye accelerates rapidly towards object. Following this acceleration, the eye’s angular velocity tends to come up with that of the object (Rashbass 1961). In human, SP gain (G-parameter, the ratio of the angular eye velocity to object velocity) usually is close to unity, i.e. eye velocity almost perfectly matches target velocity. If the object moves too fast, the eye fails to catch up with the object using the SP mechanism and turns on saccadic eye movement, making the moving object no longer acute. Similar situation arises if angular eye velocity is less than object velocity.

Nowadays, the world scientific community shows a steadily growing interest in intelligent robotics. This primarily concerns the development of humaniform robots designed to exist in the human habitat and to substitute men in the operations people used to perform themselves. Such robots should use special methods for processing information flows, adopted from the men. For this purpose, in particular, bionic SP models should be developed. Some pioneer models of the smooth pursuit eye motion, capable of explaining the experimental data, and their bionic embodiments have been already published (Shibata and Schaal 2000; Shibata et al. 2005).

All models agree upon the assumption that the SP is a negative feedback system It has been traditionally described as a servo system initiated by a mismatch between two successive stimuli image formed on the retina or their projections in brain cortex. When such a mismatch appears, the oculomotor system is commanded to smoothly compensate it (Robinson et al. 1986). However, a number of experiments demonstrate that the operation of the human SP strongly differs from that of technical servo systems. For instance, human subjects are able to continue pursuiting an object that it transiently disappears. This phenomenon was referred to as predictive pursuit (Beaker and Fuchs 1985; Madelain and Krauzlis 2003; Bennett and Barnes 2004).

Furthermore, it has been shown that the human SP is under adaptive control, which allows keeping the oculomotor response accurate in the case when the normal connection between the visual input and the motorial output is broken (Optican et al. 1985; Madelain and Krauzlis 2003; Bennett and Barnes 2004). Such observations have prompted the assumption that the eye motion is pre-programmed. A number of authors suppose that after detecting a moving object the SP system determines the direction and velocity of its motion, picks out an appropriate SP program from the set stored in memory, and finally executes this program using the oculomotor mechanism (Mitrani and Dimitrov 1978). A key element of such a system is a G-parameter

Proceedings 21st European Conference on Modelling and Simulation
Ivan Zelinka, Zuzana Oplatková, Alessandra Orsoni ©ECMS 2007
ISBN 978-0-9553018-2-7 / ISBN 978-0-9553018-3-4 (CD)

regulator (Madelain and Krauzlis, 2003; Bennett and Barnes, 2004).

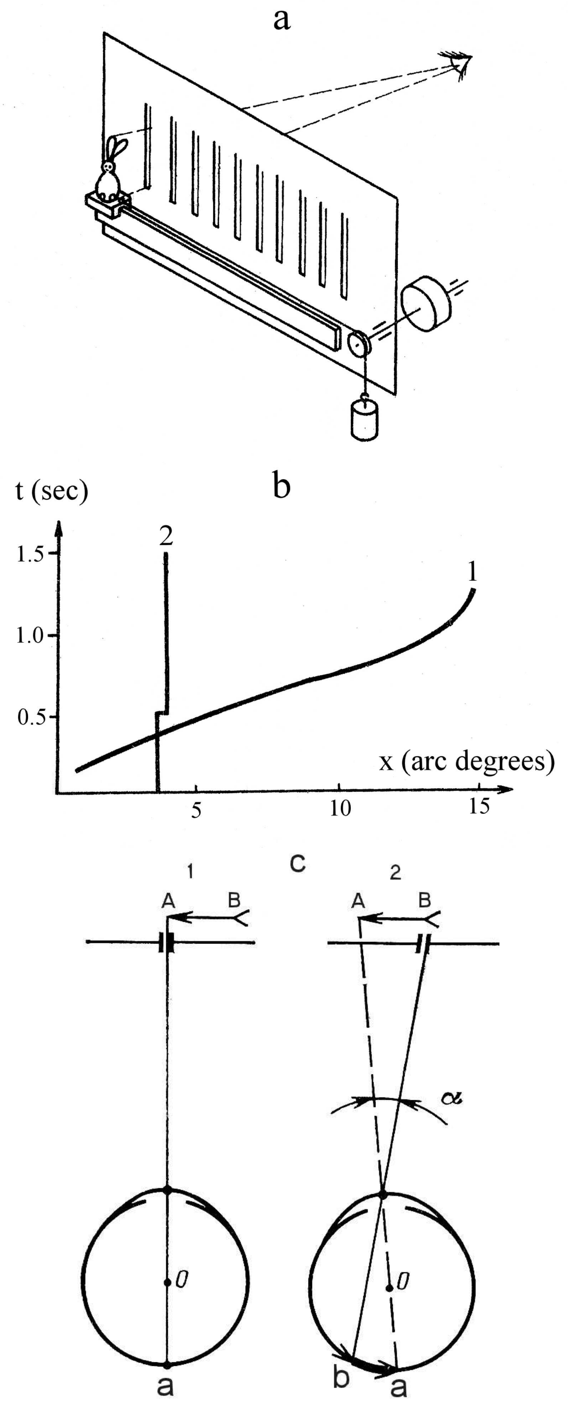

Figure 1: Psychophysical Test Scheme

Everyday human practice needs an actively controlled SP system. So, the ability to maintain the pursuit during transient disappearance of the object is useful in a natural environment where moving objects disappear behind occluding objects. However, it is still not quite clear what kind of an SP system is actually used in human practice. To clarify this point, we performed two measurement series. During the first series, the subjects were asked to track a real object that moved horizontally behind an opaque screen with a sequence of narrow vertical slits. Should a subject be able to accurately pursue the object and recognize it, this will mean that he uses an SP mechanism that differs from the simple mechanism of minimizing the retinal slip by eye movements. In the second series, we examined a large group of subjects with respect to their ability to perform the SP in the case of "apparent motion". Our primary goal was to compare the SP quality shown by different subjects and to analyze the features discovered.

SP AND MULTIPLE SHIELDING

Group of 10 subjects was asked to pursue the moving object viewed through a sequence of narrow vertical slots such that only the fragments of the object could be visible at the same time (Efimov et al. 1995). Fig.1a shows the experimental scheme of stimulus presentation. The object moved with the velocity of 13-15 deg/sec. In Fig.1a the slot width and the distance between the slots correspond to the fact. The relative sizes of the object did not exceed the distance between the slots. We presented to the subjects different objects and registered their response. A number of the subjects were recorded regarding the movements of their eyes while they observed the object. Most subjects have recognized the object after tracking it. The records of their eye movements have demonstrated that they moved smoothly with the near-constant velocity (Fig. 1b, 1). When a subject failed to recognize the object, the eye movement style was different: the subject did not smoothly pursue the object, but periodically focused his eyes on a slit (Fig. 1b, 2).

These experiments can be only explained by the fact that when tracking a moving object, the human eye and brain track rather the dynamic characteristics of the object than its current position and program the movement of the eyes on the basis of these characteristics. The programmed movement of the eyes with the tracking velocity coinciding with the object velocity provides the observer with the full retinal projection compiled from moving object's fragments perceived in series and, as a result, allows him/her to recognize the object. Fig. 1c, 1, and 2 show the schemes of forming the retinal projections of point A of the object AB at the moment t_1 (1) and its point B at the moment t_2 (2). Due to the inertia of vision, the projections of all the points of a moving object simultaneously exist on the retina for a while. As a result, at the moment t_2 the projection of the object AB forms the ab arc on the retina.

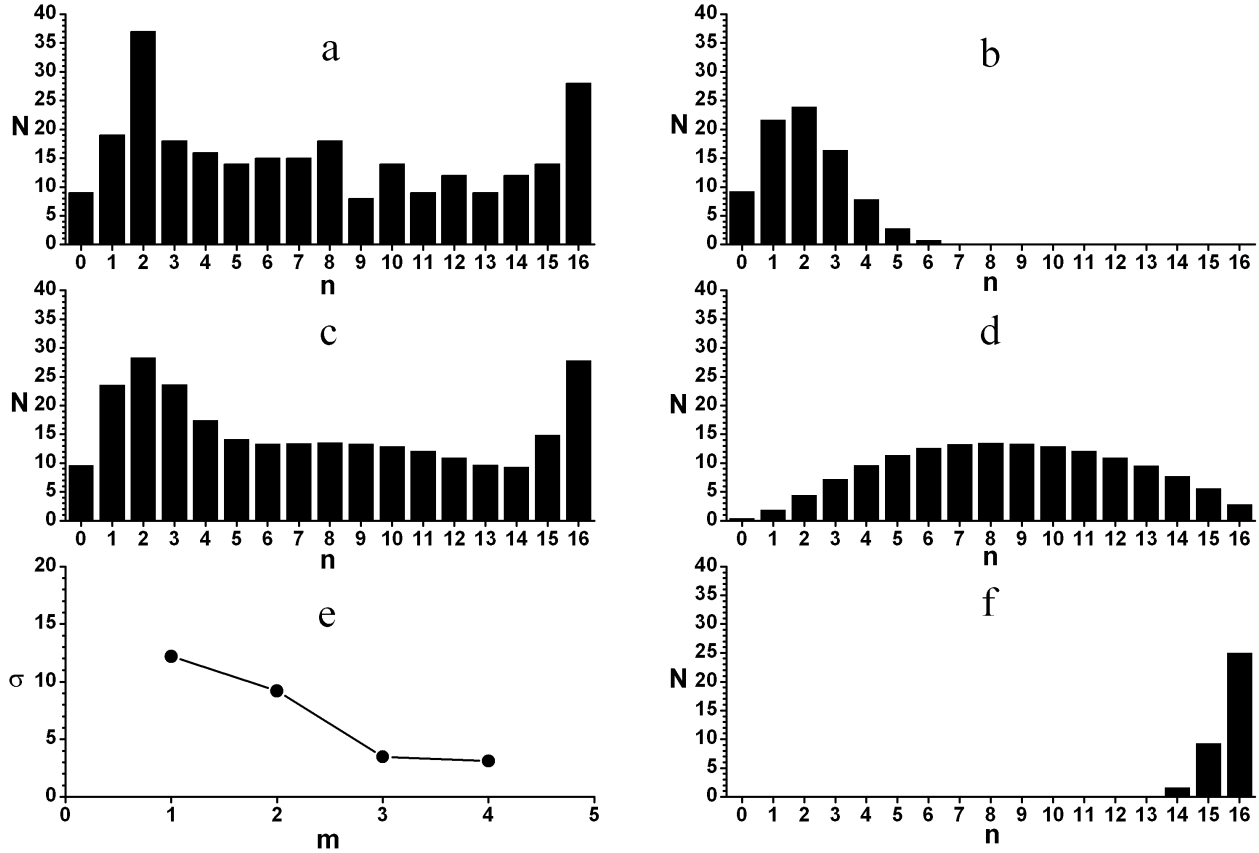

Figure 2: Decomposition of the SP Quality Histogram

SP AND APPARENT MOTION

Distinctive features of the visual perception and neural processing are not infrequently analysed in experiments producing an “apparent motion” effect. Experiments show that in the case of apparent motion the SP program is also initialized. Psychophysical experiments on the SP often use to produce the apparent vision effect a sequence of light flashes creating the impression of motion. The quality of the apparent motion depends on the temporal (Δt) and spatial (Δx) intervals between the flashes at given apparent velocity V_a defined as Δx / Δt. With increasing Δt and Δx, the percept of flash movement changes to the percept of a sequence of light flashes. Psychophysical and physiological experiments show that the apparent motion is perceived as continuous motion if Δx is less than 1° and Δt is less than 100 ms at the apparent motion velocities ranging from 10 to 20°/s (Churchland and Lisberger 2000).

Conventional methods for estimating the SP quality require a rather complex system for tracking the eye movement and therefore complicate accumulating extensive statistical data. This drawback is not peculiar to a non-contact method for estimating the SP (Surovicheva and Lebedev 1996) based on the apparent vision effect. In this method, the impression of smooth object motion is created by a sequence of flashes of LEDs separated by a fixed interflash interval, Δx about 1°. The flash is chosen long enough to modulate the LED brightness as a function of time (in the same way for all the LEDs). It allows to create illusory object as the retinal projection. If the eye makes the SP and the flash is invariable and long enough, a dash-like retinal projection is formed. On the contrary, a dot-like projection arises when the flash is short. By choosing the profile of the brightness-vs-time function of a LED, it is possible to control the shape of the apparent object as the combination of dashs and dots. Since the corresponding retinal projection is formed only as a result of the SP process, the correct observer’s answer on the object appearance is an objective measure of his/her SP ability.

We employed this rather simple device to statistically analyze the SP ability over a large group of subjects. In total, 67 young people (students of Moscow colleges and universities) have been tested. The stimuli were introduced horizontally, both from left to right and from right to left, with the velocity V_a equal to either 11°/s or 17°/s. In some cases, a wider range of V_a was used. Eight combinations of light dots and dashes were used as the stimuli. Each test series comprised 16 launches of such patterns.

The overall results of these experiments are shown in Fig.2a as a histogram of correct answers per series (n). It can be seen from the histogram that the subjects strongly differed from one another in the SP quality. Nevertheless, most of them retained the ability to program the SP in the case of apparent motion with long interstimulus intervals.

We suppose that the SP programming is a result of an innate neural mechanism, and the SP quality of a single individual depends on. As our experiments have demonstrated, the histogram of the correct answers is not a unimodal distribution. We have tested the assumption that there are m different groups of subjects, each of these described by a model distribution of the answer correctness, which characterizes the whole group. The goal of this analysis was to estimate the number of the groups and to determine the parameters of their distribution.

The ability of a subject to perform the SP was characterized by the probability p of recognizing an object moving at a given velocity. For a single test, the probability of the correct answer is

$$p'(p) = 1/8 + 7 \cdot p/8 \qquad (1)$$

(here, $1/8$ is the probability of an accidentally correct answer).

In this model, the results of multiple tests with the same subject will be represented by a histogram described by a binomial distribution:

$$N(k) \approx N_0 \cdot C_{16}^k \cdot p'^k \cdot (1-p')^{16-k},\ 0 \le k \le 16 \qquad (2)$$

where $N(k)$ is the number of series for which the subject earned k points, N_0 is the total number of the series carried out.

To describe a group of people showing close SP abilities, we chose the β-distribution (a finite unimodal distribution with controllable average and dispersion, the carrier being $[0,1]$). It is defined by the following expression:

$$f(p) = \frac{1}{A} \cdot p^\alpha \cdot (1-p)^\beta,\ A = \int_0^1 p^\alpha \cdot (1-p)^\beta \cdot dp, \qquad (3)$$

where the parameters $\alpha \ge 1$ and $\beta \ge 1$ can be expressed through the average μ and the dispersion σ^2 as follows:

$$\alpha = \mu^2 \cdot (1-\mu)/\sigma^2 - \mu,\ \beta = (1-\mu)^2 \cdot \mu/\sigma^2 - (1-\mu). (4)$$

Assuming that there are m groups of subjects, each of those described by Equation (3), we obtain the following estimate for the results of the experiment:

$$N(k) = C_{16}^k \cdot \sum_{i=1}^m N_i \cdot \int_0^1 p_0(k) dp,$$
$$p_0(k) = p'(f(p, \mu_i, \sigma_i))^k \cdot (1 - p'(f(p, \mu_i, \sigma_i)))^{16-k} \qquad , (5)$$

where (μ_i, σ_i, N_i) are the parameters of the i-th group (the average probability of recognition for this group, the RMS deviation of the recognition probability from the average, and the number of subjects in the group, respectively), and $0 \le k \le 16$.

For m ranging from 1 to 6, the parameters of the model were determined with the "shooting" method. In the course of "shooting", we minimized the RMS deviation of the model distribution (5) from the experimental histogram. The result of "shooting" for each m was the optimum (for the given number of groups) set of $3m$ parameters. As the numeric experiment has shown, a 3-group model fits the experimental data quite well (Fig. 2e), and further increasing m yields almost no improvement. Fig. 2b,d,f shows the model distributions for each found group at m=3 and fig. 2c – the total model distribution. For different groups, the SP quality strongly differed. For the first-rank group (21% of the tested subjects), the correct answer probability was 0.96±0.12, for the second-rank group (48% of the tested subjects) the probability was 0.40±0.12, for the third-rank group (31% of the tested subjects) – 0.010±0,003.

Thus, these experiments confirm that many people are able to perform the SP and recognize the object under very complex conditions, even those of the final experiment, when the object was insubstantial and repeatedly disappeared during the observation. However, such an ability is not equally inherent to all people. This is indicated by the fact that the subjects' SP characteristics appeared to be stable in repeated experiments. 33 of the 67 students were examined 3 times, with the intervals of 6 months and 1 year, respectively. During the repeated examinations, the average SP grades of the subjects did not change significantly. There was only one case when different examinations attributed the same subject to different ranks.

CONCLUSION

To explore human SP mechanisms, experiments of two types were performed with objects whose motion was interrupted in different ways. The eyes of most people tested were able to organize the SP with the speed close to that of the object motion. The obtained result can not be interpreted otherwise than a pre-programmed movement of the subject's eyes. The initial stage of the SP programming, apparently, is estimating the dynamic characteristics of the perceived object, especially its current velocity. The input parameter for this estimation is the mismatch between two successive stimuli image formed on the retina or their projections in brain cortex. On the bases of such evaluations, the algorithm predicts the optimal velocity of eye movement, required to successfully perform the SP procedure. The key feature of this prediction is the fact that it still works even if the object from time to time disappears from the field of view for a short interval. At the same time, it cannot be excluded that under simpler conditions the human visual system acts like a servo system (Robinson et al. 1986). It is also possible that the oculomotor systems of a large number of people quite rarely use mechanisms other than the servo control.

REFERENCES

Beaker, W. and A. Fuchs. 1985. "Prediction in the oculomotor system: smooth pursuit during the transient disappearance of a visual target". *Exp. Brain Res*. 57, 562-575.

Bennett, S. and G. Barnes. 2004. "Predictive smooth ocular pursuit during the transient disappearance of a visual target". *J. Neurophys*. 92, 578-590.

Churchland, M. and S. Lisberger. 2000. "Apparent motion produces multiple deficits in visually quided smooth pursuit eye movements in monkeys." *J. Neurophysiol*. 84, 216-235.

Efimov, S.K., D.G. Lebedev and N.S. Surovicheva. 1995. "Method for estimation of smooth tracing eye movements and device for it realization." *Patent RU 2033743 C1 30*, Russia.

Madelain, L. and R. Krauzlis. 2003. "Effects of learning on smooth pursuit during transient disappearance of a visual target". *J. Neurophys*., 90, 972-982.

Mitrani, L. and G. Dimitrov. 1978. "Pursuit eye movements of a disappearing moving target". *Vision Res.*, 18, 537-539.

Optican, L., D. Zee and F. Chu. 1985. "Adaptive response to ocular muscle weakness in human pursuit and saccadic eye movements". *J. Neurophys*., 54,. 110-122.

Rashbass, C. 1961. "The relationship between saccadic and smooth tracking eye movements". *J. Physiol*., 159, 326-338.

Robinson, D., J. Gordon and S. Gordon. 1986. "A model of the smooth pursuit eye movement system". *Biol. Cybern*., 55, 43-57.

Shibata, T. and S. Schaal. 2001. "Biomimetic smooth pursuit based on fast learning of the targetdynamics." *In Proceedings of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems* (Maui, Hawaii, USA). 1, 278-285.

Shibata, T.; H. Tabata; S. Schaal and M. Kawato. 2005. "A model of smooth pursuit based on learning of the target dynamics using only retinal signals." *Neural Networks*, 18, 213-225.

Surovicheva, N.S. and D.G. Lebedev 1996. "Method for estimation of smooth tracing eye movements." *Sensory systems*, 10, 124-128 (in russian).

Yarbus, A. 1967. *Eye movements and vision*. Plenum press, N.Y.

AUTHOR BIOGRAPHIES

VALENTINA P. BOZHKOVA obtained her Master degree in physics in 1966 from Moscow State University, Ph.D. degree in biology in 1972, and the degree of Doctor of Sciences (in biology) in 1987. She is with the Institute for Information Transmission Problems, Russian Academy of Sciences since 1975. Her primary research areas are biological processes modelling and psychophysiology of vision.

NADEZHDA S. SUROVICHEVA obtained her Ph.D. degree in biology in 1999. She is with the Institute for Information Transmission Problems, Russian Academy of Sciences since 1981. Her primary research areas are biological processes modelling and psychophysiology of vision.

DMITRY P. NIKOLAEV was born in Moscow, Russia. He studied physics, obtained his Master degree in 2000 and Ph.D. degree in 2004 from Moscow State University. Since 2000 he has been a research scientist at the Institute for Information Transmission Problems, RAS. His research activities are in the areas of computer vision with primary application to colour image understanding. His e-mail address is dimonstr@iitp.ru and his Web-page can be found at http://www.ddt.ru/chaddt/Lace/pub.html

DMITRY G. LEBEDEV obtained his Ph.D. degree in engineering in 1966 and the degree of Doctor of Sciences (in physics and mathematics) in 1990. Since 1966, he has been a research scientist at the Institute for Information Transmission Problems, Russian Academy of Sciences, where he is now leading a research group in the field of human vision processing and robotics.