Surgical Innovation
1–6

DOI: 10.1177/1553350615573581
sri.sagepub.com
SAGE

Three-Dimensional Eye Tracking in a Surgical Scenario

Rositsa Bogdanova, MD[1], Pierre Boulanger, PhD[1], and Bin Zheng, MD, PhD[1]

Abstract

Introduction: Eye tracking has been widely used in studying the eye behavior of surgeons in the past decade. Most eye-tracking data are reported in a 2-dimensional (2D) fashion, and data for describing surgeons' behaviors on stereoperception are often missed. With the introduction of stereoscopes in laparoscopic procedures, there is an increasing need for studying the depth perception of surgeons under 3D image-guided surgery. **Methods:** We developed a new algorithm for the computation of convergence points in stereovision by measuring surgeons' interpupillary distance, the distance to the view target, and the difference between gaze locations of the 2 eyes. To test the feasibility of our new algorithm, we recruited 10 individuals to watch stereograms using binocular disparity and asked them to develop stereoperception using a cross-eyed viewing technique. Participants' eye motions were recorded by the Tobii eye tracker while they performed the trials. Convergence points between normal and stereo-viewing conditions were computed using the developed algorithm. **Results:** All 10 participants were able to develop stereovision after a short period of training. During stereovision, participants' eye convergence points were 14 ± 1 cm in front of their eyes, which was significantly closer than the convergence points under the normal viewing condition (77 ± 20 cm). **Conclusion:** By applying our method of calculating convergence points using eye tracking, we were able to elicit the eye movement patterns of human operators between the normal and stereovision conditions. Knowledge from this study can be applied to the design of surgical visual systems, with the goal of improving surgical performance and patient safety.

Keywords

image-guided surgery, simulation, ergonomics and/or human factors study, surgical education

Introduction

Eye tracking has been frequently used in the surgical environment, either for investigating the scanning pattern of surgeons in detecting nodules in the lung and the breast[1-3] or for revealing eye-hand coordination differences between novice and expert surgeons in performing laparoscopic and endoscopic procedures.[4-7]The use of eye tracking for assessing surgeon's skills has been further studied, and it offers an objective measurement of surgical expertise.[89] Scientists in the field have introduced gaze training for improving laparoscopic technical skills in novices.[10-12]Our research group also uses eye tracking to investigate the vigilance of surgeons in the operating room.[13,14] Eye-tracking data reported in the above studies is 2-dimensional (2D), meaning only eye motions over the horizontal (*x*) and vertical (*y*) axes were measured. 2D analysis on eye-tracking data did not lead to any problem to date because most laparoscopic and endoscopic images are 2D. However, the eye movements along the longitudinal (*z*) axis are seldom mentioned.

The need for examining 3D eye motion patterns in surgeons increases with the introduction of 3D-imaging technology into the operating room. It has been claimed that 3D imaging can improve depth perception in robotic surgery. The effect of 3D vision on efficacy and accuracy in performing complex tasks on robotic systems among novices and experienced surgeons has been shown to be promising.[15,16] It was also suggested that the newly developed 3D-imaging technologies facilitated the skills development of novices in learning complex laparoscopic surgeries.[17,18] As more and more stereoscopes and stereo-images are introduced into the operating room, we believe that it is important to develop technology for tracking the

[1]University of Alberta,Edmonton, AB, Canada

Corresponding Author:
Bin Zheng, Surgical Simulation Research Lab, Department of Surgery, University of Alberta, 8440 112 St NW, Edmonton, AB Alberta T6G 2E1, Canada.
Email: bin.zheng@ualberta.ca

surgeon's eye motions in a 3D fashion. We also believe that 3D eye tracking can be further extended to other video-mediated health care procedures and image interpretation, such as those in invasive and noninvasive radiology.[19,20]

Depth perception can be defined as our ability to see and place objects at different distances from us. To perceive depth, stereoscopic vision, which is based on binocular disparity, is needed.[21] It is described by the positional difference between the 2 retinal projections to a fixated point in space. This positional difference of both images on the retina comes from the horizontal separation of our eyes, where each eye sees the world from a slightly different angle. The central visual cortex processes the different but correlated images from both eyes and computes the difference as the depth of view of the world observed. When the eyes are actively taking information from the world, the eyeballs rotate inwards (converge) and outwards (diverge) to allow the 2 optical axes to meet in front of the observer at different depth planes. The point where 2 optical axes meet is called the convergence point. The convergence point in the natural viewing condition constantly changes as the human operator needs to collect information on a target for estimating its relative distance, movement direction, speed, and so on. Knowing the location of the convergence point and its pattern of change provides us with a chance to investigate how surgeons reconstruct stereovision on the surgical site under 3D visual condition. In laparoscopic surgery, as well as in other types of video-mediated surgical procedures, the ability to reconstruct (build) a 3D model of the surgical site can be an important human factor in surgical performance and patient safety.

It is interesting that many studies have adopted 2D stereoscopic displays (stereograms) to simulate depth and study human stereovision.[22-24] Although the magnitude of binocular disparity in stereograms can vary greatly compared with real-world stereoperception, the underlying mechanism for creating stereoperception to a human operator is fundamentally the same. In viewing stereograms using the cross-eyed technique, the lines-of-sight of the observer's eyes cross in front of the image. When cross-viewed, the muscles around the eyes that control the eye lenses contract, and the image comes into focus on the fovea on the retina. The process of changing the focal length of the 2 eye lenses and placing the focus images onto the retinas is known as accommodation. Accommodation is a strong depth cue because both the muscular action and the lack of focus on different depth planes provide information on depth. In addition, in order to see close objects and fuse the images on both retinas, the observer has to converge the optical axis of both eyes on the object. As the observer is presented with a stereogram consisting of 2 side-by-side identical pictures (taken from slightly different angles), the picture on the left is to be viewed by the right eye and the picture on the right by the left eye. Creating stereoperception this way requires time for practice, which can vary between individuals.

In 1985, Cormack and Fox in New Mexico made an attempt to calculate the convergent point (also known as distance to stereoscopic form) in stereograms.[25] The reported distance to stereogram is an indirect way of computing the convergence point, which might lead to some misunderstanding. In this study, we developed a novel approach for computing the convergent point of human operators using the Tobii eye tracker. Our participants were asked to view a stereogram that consists of 2 pictures side by side by using the cross-eyed viewing technique. Once participants created stereovision, we were able to apply the new algorithm to compute the convergent point directly based on the Tobii eye-tracker data output.

Method

Participant and Experimental Environment

A controlled laboratory study was conducted in the Surgical Simulation Research Lab at the University of Alberta. A total of 12 university students with normal or corrected to normal vision participated in the study. Data from 2 of our participants were deleted because of the low quality of eye-tracking data (valid eye tracking below 50% of total task time). Thus, the data from 10 valid participants (6 men and 4 women; age = 23-48 years; mean = 32 years) were included for further analysis. Ethics approval was obtained from the Health Research Ethics Board of the University of Alberta before the recruitment of human participants. Written consent was obtained from each participant prior to entering the study.

The experimental setup included 3 main components: (1) a PC monitor with a visual stimulus consisting of 2 MRI images of the human skull base placed side by side with binocular disparity; (2) a Tobii X2 60-Hz eye tracker placed under the monitor to unobtrusively record the participant's eye motions; and (3) a chin rest mounted at 60 cm from the plane of the screen.

Task

Two MRI images were displayed on a 24-inch high-definition PC monitor (LG-24MA31D, LG Electronics, Seoul, South Korea). Sitting comfortably in front of the display, participants in the study were required to watch these 2 MRI images and create stereovision using the cross-eyed technique. The chin rest prevented any unnecessary head movements, which allowed us to record clean eye data. Once the participants were able to see the stereoimage, they maintained stereovision for 1 minute before fading back to normal vision for another minute. The trial was repeated 3 times for each participant.

Eye Tracker

A high-resolution remote eye tracker, Tobii X2 (Tobii Technology, Danderyd, Sweden) was used to capture participants' eye movements while performing picture observation. The Tobii X2 records eye motions on a monitor at 60 Hz and is accurate to within 0.5° of the visual field. With the X2's remote tracking ability, the performer's eye motions can be monitored unobtrusively while sitting at a comfortable viewing distance. Participants do not need to wear special goggles. Combined with Tobii's Studio eye-tracking analysis software suite, high-definition 3D eye motion data can be obtained for further analysis.

Variables and Algorithm

To compute convergence points, we took Tobii gaze data for the horizontal position from each eye (X1 and X2). When measuring gaze horizontal location over the screen, we set an arbitrary coordinate system, where 0 was at the left edge of the screen, and the maximum value of the horizontal location was at the right edge. For any time frame, Tobii eye-tracking output can report X_1 and X_2, and the distance between 2 gaze locations (M) was computed as $M = X_2 - X_1$, where the left eye horizontal position is subtracted from the right eye's x-axis position. Besides gaze location on the horizontal plane, we also measured observer's physical distance to the screen (D) and the interpupillary distance (E). With these 3 variables in hand, we can calculate the convergence point (F). Specifically, as displayed in Figure 1, geometrically, $D/Y = F/(E/2)$, where $Y = M - [(M - E)/2] = (M + E)/2$. Therefore, $F = (E \times D)/(E + M)$.

Convergence point measurements were compared between normal and the stereovision conditions using the SPSS (SPSS Inc, Chicago, IL) within-subject ANOVA model, and $P < .05$ was considered statistically significant.

Results

All 10 participants were able to see the stereoimage in the stereogram after a short training period on the cross-eye viewing technique. By taking their left and right eye gaze points along the x-axis and applying them to the algorithm, we were able to compute the convergence points and gaze trajectories for all participants. The convergence point between the 2 eyes was significantly closer to the participants (14 ± 1 cm) during stereovision than during normal vision (77 ± 20 cm; $P < .001$; $\eta_p^2 = 0.96$). Figure 2 displays the outcome for participant 9 over 3 cycles of task performance.

In addition to the above, we found that convergence points were distributed widely while the participants watched the 2 pictures on the screen under normal visual conditions, whereas convergence points were more centralized (or less widely distributed) when stereovision was applied (see Figure 3). Real-time data are presented in Figure 3.

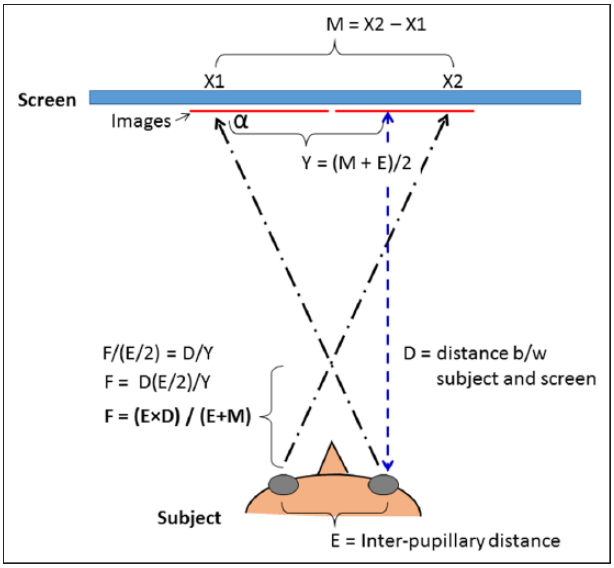

Figure 1. A geometrical representation for calculating the convergence point while a participant is viewing stereograms using the cross-eyed technique.

Discussion

With the Tobii eye tracker, we successfully computed convergence points and gaze trajectories in the surgical training environment. To compute the eye convergence point, a number of key variables are needed. These are the interpupillary distance, the distance to the stereogram plane (PC monitor), and the distance between gaze locations along the x-axis over the target. The first 2 variables can be easily obtained before the operation, and the last variable can be acquired from Tobii eye-tracker outputs at the end of the operation. By applying our algorithm using these variables, we were able to quickly calculate the convergence points for surgeons and examine their eye motion behaviors in obtaining depth information in surgical images.

The difference in convergence point distribution represents the specific gaze trajectories in the 2 visual modes (stereovision and normal vision). In detail, the eyes perform an initial transient divergent (outward) movement

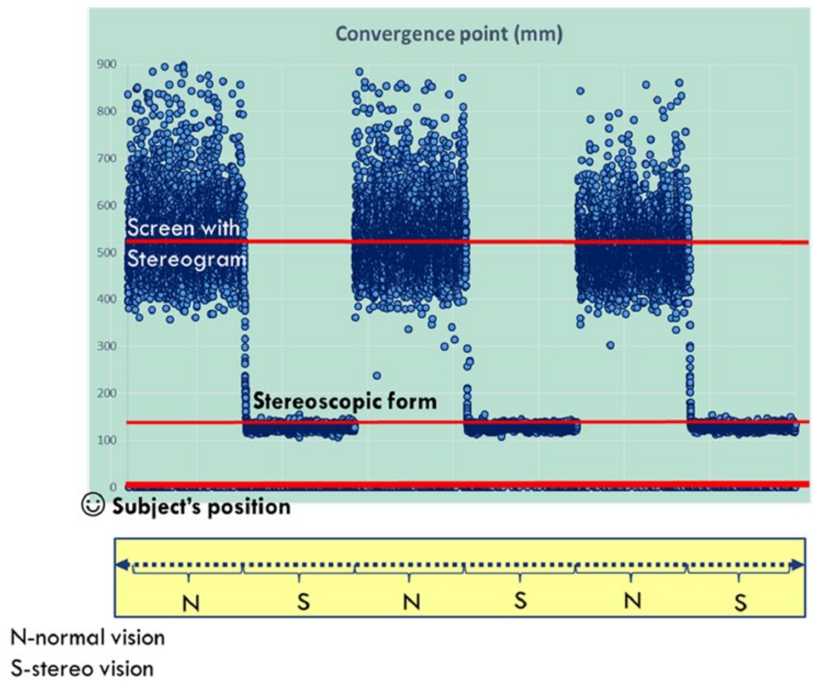

Figure 2. Convergence points of participant 9 are presented over 3 cycles between N (normal) and S (stereo) vision conditions. Note that the pictures with binocular disparity were displayed on the screen (top red line) placed at 60 cm from the observer (bottom red line).

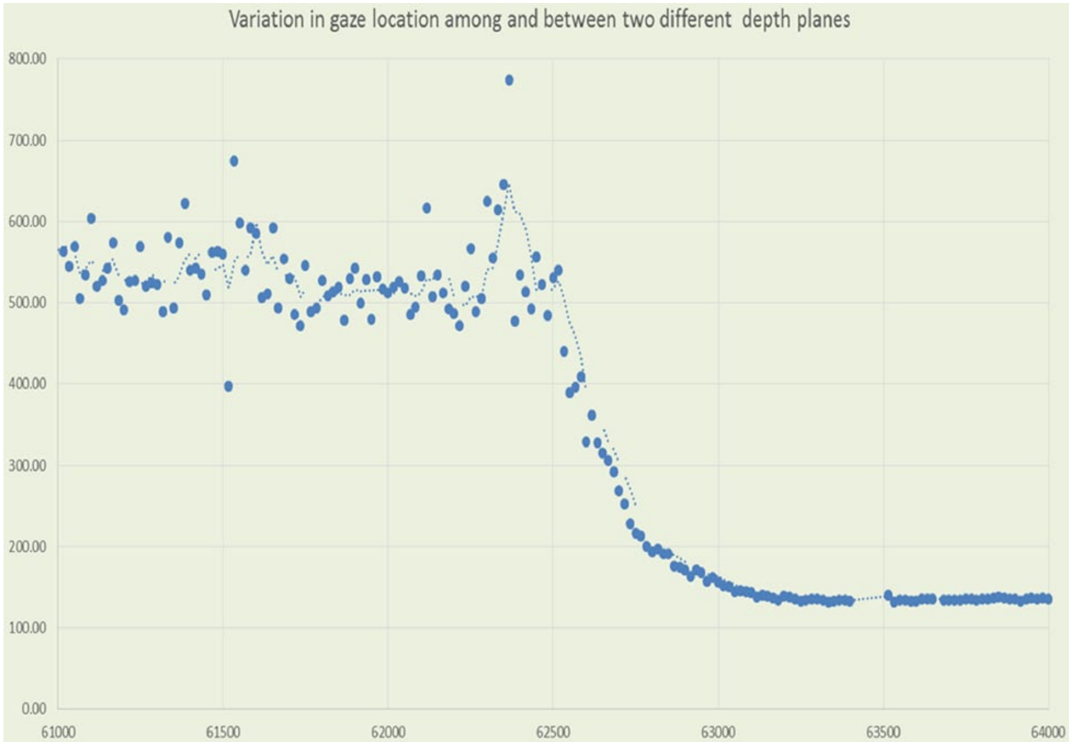

Figure 3. Moving average horizontal trajectory of the gaze at the transition period between normal (first half of the data) to stereovision (second half of the data).

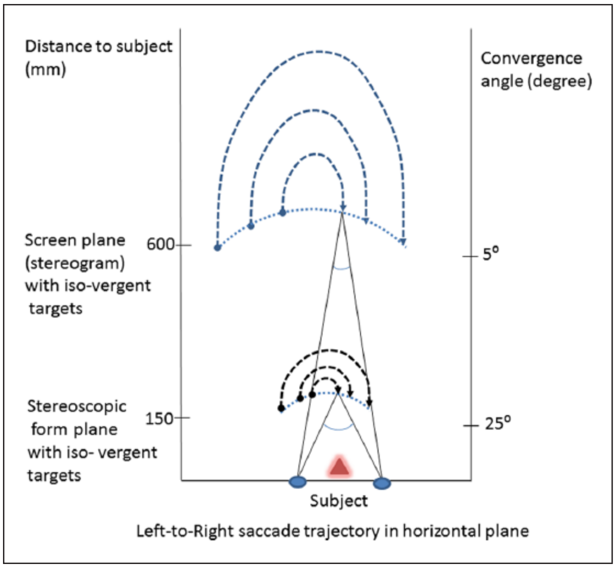

Figure 4. Intraplanar saccade trajectories between targets in 2 different depth planes (screen plane and stereoscopic form plane). The convergence angle corresponding to each depth plane is also given.

followed by the convergent (inward) movement to the nearby region of the chosen fixation point. This outward loop-like trajectory is registered and presented by eye trackers as a variation of convergence point locations at each plane of the 2 visual modes. Collewijn and Steinman[26] found that when gaze shifts between iso-vergent targets in a single depth plane, the gaze trajectory is mainly affected by the eye movement amplitude.[26] This phenomenon was recorded in the current study in a similar way. A schematic presentation of this behavior is given in Figure 4.

It is apparent that the depth of the gaze loop increases with increase in the amplitude of the horizontal gaze shift. The gaze outward loop is also affected by the convergence angle at which the eye axes meet (which presents the shift of eye gaze between different depth planes). As the convergence angle decreases with increasing distance to visual targets, a wider variation of convergence points is registered when participants look at the screen level (where the stereogram is placed). In contrast, when the eye gaze shifts to a closer fixation distance to perceive the stereoform, the variation in convergence points (also known as the depth of the outward loop) decreases significantly.

With the increasing use of eye tracking in surgery for the purposes of 3D eye movement analysis, one should be familiar with the specifics of the visual mode gaze trajectories. Convergence point calculation should be considered as an integral part of the complete 3D eye examination under 3D visual conditions. The change in convergence point between visual modes together with the subjective experience of depth when approaching 3D imaging could be used to describe the impact of 3D technology on human performance. Knowledge learned from this project could explain the visual comfort of human operators when working under reconstructed stereovision.

Conclusion

In this study, we presented a formula to calculate the convergence point using an eye-tracking system. Knowing convergence points and their patterns of change provides us with a chance to investigate how surgeons perceive stereoperception on the surgical site. While it is true that surgical tasks under the guidance of a stereoscope will be much more complicated than the simulation setting reported in this article, the computing algorithm developed in this project is an initial step toward our goal of studying stereoperception of surgeons in 3D image-guided procedures. We will further examine this issue with a wider range of surgical tasks. The ability to create correct depth perception can be an important human factor in surgical performance and patient safety. In this sense, knowledge gained from this project will help us understand surgeons' behaviors and design a better surgical visual system for image-guided procedures.

Author Contributions

Concept and design: Bogdanova, Zheng, Boulanger
Data collection, analysis and interpretation: Bogdanova, Zheng
Manuscript preparation: Bogdanova, Zheng
Critical comments on manuscript: Boulanger

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: This project was partially supported by Edmonton Civic Employees Research Awards 2013 to Dr Bogdanova and Dr Zheng for studying of stereoperception of surgeons in performing laparoscopic surgery.

References

1. Kundel HL, Nodine CF. A visual concept shapes image perception. *Radiology*. 1983;146:363-368.
2. Nodine CF, Kundel HL, Lauver SC, et al. Nature of expertise in searching mammograms for breast masses. *Acad Radiol*. 1996;3:1000-1006.
3. Nodine CF, Mello-Thoms C, Kundel HL, et al. Time course of perception and decision making during mammographic interpretation. *AJR Am J Roentgenol*. 2002;179:917-923.
4. Atkins MS, Jiang X, Tien G, et al. What do surgeons see: capturing and synchronizing eye gaze for surgery applications. *Surg Innov*. 2013;20:243-250.

5. Atkins MS, Nicolaou M, Yang GZ. Eye monitoring applications in medicine. In: Hammoud RI, ed. *Passive Eye Monitoring: Algorithms, Applications and Experiments*. Berlin, Germany: Springer-Verlag; 2008:323-346.
6. Krupinski EA, Tillack AA, Richter L, et al. Eye-movement study and human performance using telepathology virtual slides: implications for medical education and differences with experience. *Hum Pathol*. 2006;37:1543-1556.
7. Law B, Atkins MS, Kirkpatrick AE, et al. Eye gaze patterns differentiate novice and experts in a virtual laparoscopic surgery training environment. Paper presented at: Symposium on Eye Tracking Research & Applications; March 22-24, 2004; San Antonio, TX.
8. Nicolaou M, James A, Darzi A, Yang G-Z. *A Study of Saccade Transition for Attention Segregation and Task Strategy in Laparoscopic Surgery*. Berlin, Germany: Springer; 2004.
9. Sodergren MH, Orihuela-Espina F, Clark J, Darzi A, Yang GZ. A hidden Markov model-based analysis framework using eye-tracking data to characterise reorientation strategies in minimally invasive surgery. *Cogn Process*. 2010;11:275-283.
10. Wilson MR, Vine SJ, Bright E, Masters RS, Defriend D, McGrath JS. Gaze training enhances laparoscopic technical skill acquisition and multi-tasking performance: a randomized, controlled study. *Surg Endosc*. 2011;25:3731-3739.
11. Sodergren MH, Orihuela-Espina F, Froghi F, et al. Value of orientation training in laparoscopic cholecystectomy. *Br J Surg*. 2011;98:1437-1445.
12. Vine SJ, Masters RS, McGrath JS, Bright E, Wilson MR. Cheating experience: guiding novices to adopt the gaze strategies of experts expedites the learning of technical laparoscopic skills. *Surgery*. 2012;152:32-40.
13. Tien G, Zheng B, Atkins MS. Quantifying surgeons' vigilance during laparoscopic operations using eyegaze tracking. *Stud Health Technol Inform*. 2011;163:658-662.
14. Zheng B, Tien G, Atkins SM, et al. Surgeon's vigilance in the operating room. *Am J Surg*. 2011;201:673-677.
15. Badani KK, Bhandari A, Tewari A, Menon M. Comparison of two-dimensional and three-dimensional suturing: is there a difference in a robotic surgery setting? *J Endourol*. 2005;19:1212-1215.
16. Byrn JC, Schluender S, Divino CM, et al. Three-dimensional imaging improves surgical performance for both novice and experienced operators using the da Vinci Robot system. *Am J Surg*. 2007;193:519-522.
17. Bhayani SB, Andriole G. Three-dimensional (3D) vision: does it improve laparoscopic skills? An assessment of a 3D head-mounted visualization system. *Rev Urol*. 2005;7:211-214.
18. Votanopoulos K, Brunicardi FC, Thornby J, Bellows CF. Impact of three-dimensional vision in laparoscopic training. *World J Surg*. 2008;32:110-118.
19. Getty DJ. Clinical medical applications for stereoscopic 3D displays. *J Soc Inf Display*. 2007;15:377-384.
20. Liu X, Jiang H, Lang Y, et al. A novel stereoscopic projection display system for CT images of fractures. *Exp Ther Med*. 2013;5:1677-1682.
21. Wheatstone C. On some remarkable and hitherto unobserved phenomena of binocular vision. *Optom Wkly*. 1962;53:2311-2315.
22. Poggio GF, Motter BC, Squatrito S, et al. Responses of neurons in visual cortex (V1 and V2) of the alert macaque to dynamic random-dot stereograms. *Vision Res*. 1985;25:397-406.
23. Pressmar SB, Haase W. Effects on depth perception and pattern recognition in random dot stereograms by changing the matrix dot density [in German]. *Ophthalmologe*. 2001;98:955-959.
24. Julesz B. Binocular depth perception of computer-generated patterns. *Bell Syst Tech J*. 1960;39:1125-1162.
25. Cormack R, Fox R. The computation of disparity and depth in stereograms. *Percept Psychophys*. 1985;38:375-380.
26. Collewijn HEC, Steinman RM. Trajectories of the human binocular fixation point during conjugate and non-conjugate gaze-shifts. *Vision Res*. 1997;37:1049-1069.