UPCommons

Portal del coneixement obert de la UPC

http://upcommons.upc.edu/e-prints

Title: VISUAL MECHANISMS GOVERNING THE PERCEPTION OF AUTO-STEREOGRAMS

Running head: Stereopsis in auto-stereogram perception

Authors: Aurora Torrents Gómez, M.Sc.[*]

Núria Lupón, Ph.D.[*]

Genís Cardona, Ph.D.[*]

J. Antonio Aznar-Casanova, Ph.D.[^]

[*] Department of Optic & Optometry, Technical University of Catalonia (Spain)

[^] Department of Basic Psychology and Institute for Brain, Cognition and Behaviour (IR3C), University of Barcelona (Spain)

Corresponding author: J. Antonio Aznar Casanova

Faculty of Psychology

University of Barcelona

Passeig de la Vall d'Hebron, 171

E08035, Barcelona (Spain)

E-mail: jaznar2@ub.edu

Date: 25th February 2011, Revised 12th May 2011, Revised R1 27th June 2011, Revised R2 1st July 2011

ABSTRACT

BACKGROUND

Single Image Random Dot Stereograms (SIRDS) have been employed to study diverse visual parameters and skills. The aim of this study was to identify the main optometric factors involved in the perception of SIRDS, as well as to obtain a discriminant model to categorize our participants in terms of their skill in perceiving SIRDS.

METHODS

Response time was determined in order to assess the ability of a total of 69 participants to perceive the hidden 3D shape in an auto-stereogram presented under controlled conditions, whereupon three skill level groups were defined. The same participants were administered a battery of optometric tests to evaluate various aspects of accommodation and convergence, as well as stereopsis and phoria. Linear discriminant analysis, which served to examine the relationship between response times and the evaluated visual parameters and skills, provided a set of discriminant functions (or model), thus allowing for the categorization of participants according to their skill to perceive SIRDS.

RESULTS

Two discriminant functions were obtained, which allowed for an overall predictive accuracy of 66.67% (p = 0.024), with a higher predictive accuracy for groups 1 (78.26%) and 2 (75.86%) than for group 3 (35.29%). Stereoacuity, negative relative convergence, phoria at near and, to a lesser extent, AC/A ratio were found to be the most relevant

discriminant variables, although between-group statistically significant differences were only disclosed for stereoacuity (p = 0.001) and negative relative convergence (p = 0.003).

CONCLUSION

The ability to perceive SIRDS was found to be related to many visual parameters and skills, including, but not limited to, stereoacuity and negative relative convergence. However, it is uncertain whether SIRDS may be considered a useful tool in clinical practice.

Keywords: Accommodation and Convergence; Auto-stereogram; Binocular vision; SIRDS (Single Image Random Dot Stereogram); Stereoacuity

INTRODUCTION

One of the most amazing characteristics of the binocular visual system is the ability to integrate two two-dimensional (2D) images into a single stereoscopic (3D) image. The importance of binocular vision in guiding our interaction with the environment is shown in almost every aspect of daily life. Stereoscopic vision may be evaluated through the dichoptic presentation of stimuli as in random dot stereograms (RDS), first developed by Julesz in 1960[1-3].

A Single Image Random Dot Stereogram (SIRDS) consists in a single two-dimensional image containing almost identical horizontally repeating patterns. When viewed with the proper vergence, a hidden three-dimensional scene is observed as emerging from the plane of the stereogram[4,5]. In contrast to RDS, SIRDS do not require any external device in order to dichoptically present two images to the observer. In addition, SIRDS may be designed with larger disparities than RDS[6].

The observation of SIRDS is partly based on the well known wallpaper illusion first documented by Brewster in 1844[7], who described how a horizontally repetitive pattern appeared to shift in depth, either behind (parallel fixation) or forward (crossed fixation) with reference to the plane of the wallpaper. The observer attempts to match two consecutive patterns, which appear to originate in the same object, through the left and right eyes, thus creating a shift in the perceived depth of the object[8]. A geometrical explanation of this phenomenon in the case of parallel fixation is shown in **Figure 1**. This figure shows an observer with the plane of accommodation coincident with the plane of the target (repetitive pattern) while, at the same time, the plane of

convergence is located behind the plane of accommodation. This dissociation between accommodation and convergence serves to reveal the apparent or hidden object.

As depth perception depends on the angle of convergence, the location of the "apparent" or virtual object is determined by the distance between identical elements in the auto-stereogram. The repeating pattern in SIRDS consists in a vertical strip containing a random dot map[4,9,10], thus allowing for a multiplicity of possible distances between repeated objects (individual dots) to be erroneously matched by the observer's visual system (the well-known false matching reported by Marr in 1982[11]), resulting in many possible planes in depth, and leading to the impression of volume (see **Figure 2**).

When viewing a real scene, accommodation and convergence normally function as an integrated, coupled mechanism. For the correct visualization of SIRDS, however, a disassociation of the planes of accommodation and convergence must occur[12,13], as displayed in **Figure 1**.

It is common practice for optometrists to examine the accommodation/convergence relationship by measuring the amplitude of the zone of clear single binocular vision (ZCSBV), which informs about the ability to uncouple the planes of accommodation and convergence[14]. The ZCSBV is assessed through the conjoint evaluation of the positive and negative relative convergence (PRC and NRC), that is the amount of base-out or base-in prisms until blur, respectively, and positive and negative relative accommodation (PRA and NRA), that is, the maximum ability to stimulate or relax

accommodation while maintaining clear, single binocular vision, respectively. In addition, the relationship between accommodative convergence and accommodation (AC/A ratio)[13] and the presence of a small heterophoria may also influence the visualization of SIRDS, as well as the natural stereoacuity of the observer.

The ability to perceive of SIRDS has been investigated with reference to the visual skills of the observer[15,16], mainly as a means to develop new visual tests. Indeed, the self reported skill to visualize SIRDS has been found to be highly predictive of stereoacuity, as measured by the TNO test[17]. Other investigators have explored the association between the time required to perceive SIRDS and a variety of visual skills, noting that most observers can correctly see the hidden stereo image in less than 20 seconds provided they initiate or maintain the proper amount of divergence[18]. The inability of some observers to perceive SIRDS, even after several attempts lasting longer than one minute, has been explained by their ignorance of the proper viewing strategy, that is, their lack of information, or practice, regarding the need to disassociate accommodation and convergence[18-20].

It is interesting to note that, to the best of our knowledge, it is not clear why some observers with good stereoacuity fail to perceive SIRDS. Therefore, a study was designed aiming at identifying the visual mechanisms involved in the ability to perceive the hidden stereo image in SIRDS, as well as at obtaining a discriminant model to categorize our participants in terms of their skill in perceiving SIRDS. Participants were grouped into three different categories, according to their ability to perceive SIRDS, and the predictive value of various visual aspects was examined.

METHODS

Participants

A total of 69 healthy volunteers (15 males, 54 females) aged between 21 and 28 years (mean age 23.43 years; SD = 4.85 years) participated in the study. Participants were recruited from the student population of the Technical University of Catalonia. Only participants with monocular and binocular corrected distance and near visual acuity ≥ 1 (decimal notation) were included in the study. Exclusion criteria were manifest binocular visual imbalance, color vision anomalies, existing ocular pathology, ongoing ocular treatment and history of ocular or refractive surgery.

All participants provided written informed consent after the nature of the study was explained to them. The study was conducted in accord with the Declaration of Helsinki tenets of 1975 (as revised in Tokyo in 2004) and was approved by the institutional ethical board of the Technical University of Catalonia

Experimental setting

The same random dot auto-stereogram from the book "Magic Eye: A New Way of Looking at the World" (N. E. Thing Enterprises, 1993)[21] was employed as the target stimulus throughout the study. This auto-stereogram, which showed a repeating pattern of red roses, with a heart as the hidden 3D shape, displayed a size of 19 cm x 25.2 cm and was placed on a lectern in front of the observer. Luminance of the stimulus remained constant at 100 cd/m2.

Procedure

Once our sample was defined by adhering to the mentioned inclusion and exclusion criteria, all participants were administered a complete visual examination in terms of the visual parameters and skills assumed to govern the perception of SIRDS. This part of the study was conducted by a single experienced optometrist, who was unaware of the aims of the investigation.

The following aspects were examined: near vision interpupillary distance at 40 cm (NID), with a pupillometer (HX-400 PD Meter); stereoacuity, measured with the TNO test[22] at 40 cm; phoria in distance and near vision, measured with the cover test and a handheld prism bar; positive and negative relative accommodation; positive and negative relative convergence (measured at 40 cm); AC/A ratio, evaluated with the gradient method. These testing procedures are well described in published literature (see, for example, Borish's Clinical Refraction[23]).

Participants were then seated in front of a table provided with a head and chin rest to prevent unwanted head movements. The auto-stereogram was placed on a non-fixed lectern allowing the participants to adjust their observation distance. This lectern was initially moved to a very short distance from the observers, whereupon the participants were instructed to slowly and progressively increase the observation distance until the hidden 3D shape became visible. A time limit of 90 seconds was considered adequate to decide whether the observation of SIRDS was possible or not, and all participants were encouraged to keep trying, even allowing for observation distance to be reversed, until this time limit was reached. Once the hidden shape was

visible participants had to press a button, thus registering the minimum time required to perceive the auto-stereogram, whereupon they were asked to describe the 3D shape they saw. This time, as previously reported[24], served in the categorization of all participants in terms of their ability to perceive SIRDS.

It is interesting to mention that none of our participants had had any previous experience with SIRDS, nor was subject to any specific training prior to the testing session. In addition, participants did not receive any instructions on how to perceive the stereogram.

Data analysis

Data were analyzed by means of Linear Discriminant Analysis (LDA). This procedure is useful to classify a set of subjects into several predetermined categories or groups, according to a set of independent variables called predictors or discriminant variables (x_1, x_2, ..., x_n). The groups are determined by the values of the dependent or grouping variable. The model is based on a set of participants ("training set") with known category and known values of the discriminant variables, whereupon the LDA constructs one or more linear functions of the predictors, or discriminant functions ($f = a_0 + a_1\ x_1 + a_2\ x_2 + ... + a_n\ x_n$), which allow for the categorization of new subjects with known values of their discriminant variables. The number of discriminant functions is defined as (g-1), where g is the number of categories. As it is recommended when the sample size is relatively small[25], discriminant functions were developed for the entire sample and then employed to classify the participants of the same study group into the established categories.

Three skill levels or categories were defined, according to the minimum time (RT) that was required to perceive the hidden 3D shape in the auto-stereogram (dependent variable): group 1 ($0 < RT < 10$ s), group 2 (RT > 10 s), and group 3 (SIRDS were not perceived). We opted for 10 seconds as our cut-off value as it corresponded to the median of our study sample (time interval values were not normally distributed) (see **Figure 3**). Discriminant or predictor variables were NID, stereoacuity, phoria in distance and near vision, PRC, NRC, PRA, NRA and AC/A ratio. The Linear Discriminant Analysis was performed with the SPSS software (version 17.0, SPSS Inc., Chicago, IL, USA).

RESULTS

A summary of the discriminant variables for each skill level group is displayed in **Table 1** (as mean ± standard deviation). It may be noted that NRA values are higher than should be expected for a testing distance of 40 cm, which may lead to suggest that our study group included some low, non corrected hyperopes. Nonlinear relationships are not reflected in the discriminant functions unless specific variable transformations are made to represent nonlinear effects. This was the case with both near and distant vision phoria. Therefore, it was assumed that participants with a slight heterophoria (phoria = -1) required less time to perceive SIRDS and a new variable was defined as the absolute value of the phoria plus 1 [abs(phoria+1)].

The relative role of each discriminant variable in the categorization of our participants, as displayed in **Table 1**, revealed that, whereas stereoacuity values could discriminate between group 3 and groups 1 and 2, but failed to differentiate between group 1 and group 2, NRC and phoria at near could accurately discriminate between group 1 and

group 2 (a preliminary analysis revealed shorter RT in exophoria than in esophoria). Differences between groups, when submitted to an ANOVA analysis, were found to reach statistical significance only for stereoacuity (p = 0.001) and NRC (p = 0.003).

Table 2 summarizes the standardized coefficients of the discriminant functions f_1 and f_2 resulting from LDA. The relative values of these coefficients show that stereoacuity (-0.633 and 0.746), phoria at near (0.683 and 0.664) and NRC (0.651 and 0.424) were the most relevant predictors for the categorization of the participants in the present study. To a lesser extent, AC/A ratio (-0.232 and -0.330) could also be considered a relevant variable.

The localizations of the participants according to the values obtained from both discriminant functions are shown in **Figure 4**. The group centroid for each of the three skill level categories is also represented. The horizontal and vertical separation between group centroids reflects the higher discrimination power of f_1 in comparison to f_2.

The discriminant functions served as a model to assign a group to each participant in accordance with the values of their respective predictor variables. **Table 3** compares the actual classification of all participants with the one predicted by the model. It may be observed that, overall, 66.67% of participants were correctly categorized in one of the three skill level groups, although the discriminant functions were found to present a higher predictive accuracy for groups 1 (78.26%) and 2 (75.86%) than for group 3, in

which case only 35.29% of participants previously identified as unable to perceive SIRDS were correctly allocated through discriminant analysis of the predictor variables.

It must be noted that our statistical analysis revealed a certain overlap between groups, as shown by small eigenvalues of the discriminant functions and moderate values of the canonical correlation of the discriminant functions (**Table 4**), as well as Wilks' lambda values close to 1 (**Table 5**). However, contrast of functions disclosed a statistically significant p-value = 0.024 for the whole model (f_1 to $f2$), thus concluding that the model (including both functions) could be considered as discriminating.

DISCUSSION

The present study aimed at exploring the predictive accuracy of diverse visual parameters and skills, mainly related to accommodation and convergence, to correctly categorize observers according to their ability to perceive SIRDS. Categorization *per se* was not a goal of the present investigation. Data analysis through LDA was considered a useful tool to examine the relative contribution of each predictor to the explanation of the high intersubject variability in SIRDS perception.

The ability to perceive SIRDS was found to be described within an accuracy of 66.67% by the discriminant functions resulting from LDA. However, whereas the discriminant functions offered a high predictive accuracy for groups 1 and 2, only one third of participants previously allocated to group 3 were properly categorized. Indeed, of a total of 52 participants from groups 1 and 2 (23 + 29), only 3 observers were assigned

to group 3, that is, 94.23% of those participants capable of discovering the hidden 3D image in the auto-stereogram were correctly identified by our discriminant model.

The inability of some observers with normal stereovision to perceive SIRDS has been previously documented[18,19], suggesting that a certain amount of practice is required to achieve the desired dissociation of the automatic and natural coupling between accommodation and convergence. In addition, these authors reported higher success rates when observers were informed of the most appropriate viewing strategy to perceive SIRDS[20]. Therefore, it may be speculated whether, with the adequate information and proper training, many participants erroneously classified in group 3 would have been correctly allocated to group 1 or group 2.

Viewing distance is of critical importance for the correct visualization of SIRDS, which involves three different depth cues (binocular disparity, accommodation and convergence). Even if our results failed to disclose any significant association between RT and viewing distance, it may be interesting to mention that measurements could have been, at least to some extent, confounded by the personality of each subject (some may be prepared to respond very quickly, while others may deliberate for longer, even after they have perceived the stereo image) or by difficulty in moving the lectern (although it presented a smooth movement, some participants may have experienced difficulties in adjusting it to the proper distance at which the target could be seen).

The relative values of the standardized coefficients of the discriminant functions f_1 and f_2 revealed stereoacuity, NRC, phoria at near and, to a lesser extent, AC/A ratio as the most relevant predictors for the categorization of our participants. This finding is in agreement with the previously documented association between stereoacuity and the self-reported skill in perceiving Magic Eye™ stereograms, as measured by the standard clinical TNO test[17], and would suggest that, apart from stereoacuity, divergence is the most important factor to consider when exploring the ability to perceive SIRDS. Indeed, an ANOVA analysis of all discriminant variables disclosed statistically significant between-group differences only for stereoacuity ($p = 0.001$) and NRC ($p = 0.003$). It may be noted that, from an optometric standpoint, this finding is not unexpected, since all participants perceived SIRDS with parallel fixation, that is, uncrossed disparity.

In order to further explore the predictive value of stereoacuity and NRC, a stepwise discriminant analysis approach was implemented in which only stereoacuity and NRC were taken into consideration in the categorization process. The results of this analysis revealed that, although statistical significance was attained ($p < 0.001$), the simplified model only allowed for the correct categorization of 52.17% of participants, as compared with a predictive accuracy of 66.67% with the complete model (nine discriminant variables). Therefore, it may be concluded that, even though stereoacuity and NRC play a significant role in explaining the high intersubject variability in SIRDS perception, other visual parameters and skills are necessary to gain a better understanding of this particularly interesting visual phenomenon.

Finally, it is important to note that, although some of the differences between the numerical values of the visual parameters and skills under evaluation lead to fair accuracy in discriminating our participants according to their ability to perceive SIRDS, many of the same differences could not be considered as clinically significant. This fact may account for some of our findings being unexpected or difficult to explain. In effect, NRC values were expected to be higher in those participants requiring less time to perceive SIRDS and lowest in participants pertaining to group 3. However, the opposite trend was observed, even though there was no association between NRC and viewing distance. In addition, our analysis disclosed a moderate between-group overlap and, even if not investigated in the present study, a possible influence of training and knowledge of viewing strategies in SIRDS perception (which could partially account for those participants failing to perceive the hidden shape altogether). All these factors probably advise against the implementation of SIRDS as a new tool for practitioners to employ in their daily visual examination routine.

In conclusion, the ability to perceive SIRDS is related to many visual parameters and skills, including, but not limited to, stereoacuity and negative relative convergence. However, it is uncertain whether SIRDS may be considered a useful tool in clinical practice.

References

1. Julesz B. Binocular depth perception of computer-generated patterns. *Bell System Tech J* 1960; 39: 1125-1161.

2. Julesz B. Binocular depth perception without familiarity cues. *Science* 1964; 145: 356–363.

3. Julesz B. Foundations of Cyclopean Perception. Chicago: The University of Chicago Press, 1971.

4. Tyler CW. Sensory processing of binocular disparity. In: Schor CM, Ciuffreda KJ, eds. Vergence Eye Movements: Basic and Clinical Aspects. Boston: Butterworths, 1983. p 199–295.

5. Tyler CW, Clarke MB. The Autostereogram. *Proceedings of SPIE* 1990; 1256: 182–197.

6. Davis H, Buckley D, Jacobs REG, Brennand DAA, Frisby JP. Accommodation to large disparity stereograms. *J AAPOS* 2002; 6: 377-384.

7. Brewster D. On the knowledge of distance given by binocular vision. *T Roy Soc Edin* 1844; 15: 663-674.

8. McKee SP, Verghese P, Ma-Wyatt A, Petrov Y. The illusion wallpaper explained. *J Vis* 2007; 7: 10, 1-11.

9. Dyckman, D. Single Image Random Dot Stereograms. *Stereo World* 1990; 17: 12-13.

10. Tyler CW. The birth of computer stereograms for unaided stereovision. In: Horibuchi S, ed. Stereogram. San Francisco: Cadence Books, 1994. p 83–89.

11. Marr D. Vision. A computational investigation into the human representation and processing of visual information. New York: WH Freeman, 1982.

12. Pinker S. The Mind's Eye. In How the Mind Works. New York: Norton, 1997. p 211–233.

13. McLin LN Jr, Schor CM. Voluntary effort as a stimulus to accommodation and vergence. *Invest Ophthalmol Vis Sci* 1988; 29: 1739-1746.

14. Goss DA. Ocular Accommodation, Convergence and Fixation Disparity: a Manual of Clinical Analysis, 2nd ed. Newton, MA: Butterworth-Heinemann, 1995.

15. Reimann D, Ditzingerl T, Fischer E, Haken H. Vergence eye movement control and multivalent perception of autostereograms. *Biol Cybern* 1995; 73: 123–128.

16. Portas CM, Strange BA, Friston KJ, Dolan RJ, Frith CD. How does the brain sustain a visual percept? *Proc Biol Sci* 2000;267: 845-850.

17. Wilmer JB, Backus BT. Self-reported Magic Eye stereogram skill predicts stereoacuity. *Perception* 2008; 37: 1297-1300.

18. Lee DY, Ireland D. Single Image Random Dot Stereogram: principles and clinical implications. *J Am Optom Assoc* 1996; 67: 209-214.

19. Richardson D. Create stereograms on your PC: Discover the World of 3D Illusion. Corte Madera, CA: Waite Group Press, 1994.

20. Son TM, Marosi G, Gschwindt A. Case Study of autostereoscopic image based on SIRDS algorithm. *Periodica Polytechinca Ser El Eng* 2001; 45: 119-138.

21. N. E. Thing Enterprises. Magic Eye: A New Way of Looking at the World. Kansas City: Andrews and McMeel, 1993. p 8.

22. Walraven J. Amblyopia screening with random-dot stereograms. *Am J Ophthalmol* 1975; 80: 893-900.

23. Benjamin WJ. Borish's Clinical Refraction. 2nd ed. St Louis: Butterworth-Heinemann, 2006.

24. Bradshaw MF, Rogers BJ, De Bruyn B. Perceptual latency and complex random-dot stereograms. *Perception* 1995; 24:749 – 759.

25. Hair JF, Anderson Jr. RE, Tatham RL, Black WC. Multivariate data analysis. 5th ed. Englewood Cliffs, NJ: Prentice Hall, 1998.

TABLES

Table 1. Descriptive statistics (means and standard deviations) of all predictor values

for each group

Discriminant variables	Group 1		Group 2		Group 3	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
NID	56.32	2.75	55.51	2.27	56.20	2.64
Stereo acuity	55.00	43.09	55.86	30.23	154.41	167.51
Abs(Phoria+1)@near	2.04	1.66	3.65	3.28	3.00	3.46
Abs(Phoria+1)@far	1.95	1.49	3.24	4.28	2.47	2.47
PRA	-3.37	1.48	-3.92	1.82	-3.75	2.10
NRA	3.21	0.64	3.51	0.84	3.41	0.81
PRC	27.65	10.56	28.31	11.16	25.23	9.99
NRC	13.65	3.11	17.69	6.66	12.82	3.81
AC/A	2.30	1.26	2.37	1.26	2.67	1.97

Table 2. Standardised coefficients of the canonical discriminant functions

VARIABLES	Coefficients $f1$	Coefficients $f2$
NID	0.016	-0.174
Stereoacuity	-0.633	0.746
Abs(Phoria+1) @near	0.683	0.664
Abs(Phoria+1) @far	-0.139	-0.212
PRA	-0.056	-0.260
NRA	-0.025	0.242
PRC	0.034	-0.285
NRC	0.651	0.424
AC/A ratio	-0.232	-0.330

Table 3. Model prediction versus actual group allocation in number and percentage of participants

		Predicted group		
Actual group	**Group size**	**1**	**2**	**3**
1	23	18 (78.26%)	2 (8.70%)	3 (13.04%)
2	29	7 (24.14%)	22 (75.86%)	0 (0.00%)
3	17	7 (41.18%)	4 (23.53%)	6 (35.29%)
Overall correct categorization: **66.67%**				

Table 4. Eigenvalues, relative percentage of variance explained by each discriminant function and canonical correlation of the discriminant functions (small eigenvalues of the discriminant functions and moderate values of the canonical correlation of the discriminant functions reveal a certain overlap between groups)

Discriminant function	Eigenvalue	Relative percentage	Canonical correlation
f_1	0.386	65.65	0.528
f_2	0.202	34.35	0.410

Table 5. Wilks' lambda and p-values testing the discriminating power of our model (Wilks' lambda values close to 1 describe a small overlap between groups. A p value of 0.024 for the whole model (f_1 to f_2) reflects a high discriminating power)

Contrast of functions	Wilks' lambda	p-value
f_1 to f_2	0.600	0.024
f_2	0.832	0.180

FIGURES

Figure 1. The observation of a horizontally repeating pattern may lead, in uncrossed disparity, to the perception of an "apparent" or virtual object located behind the plane of accommodation.
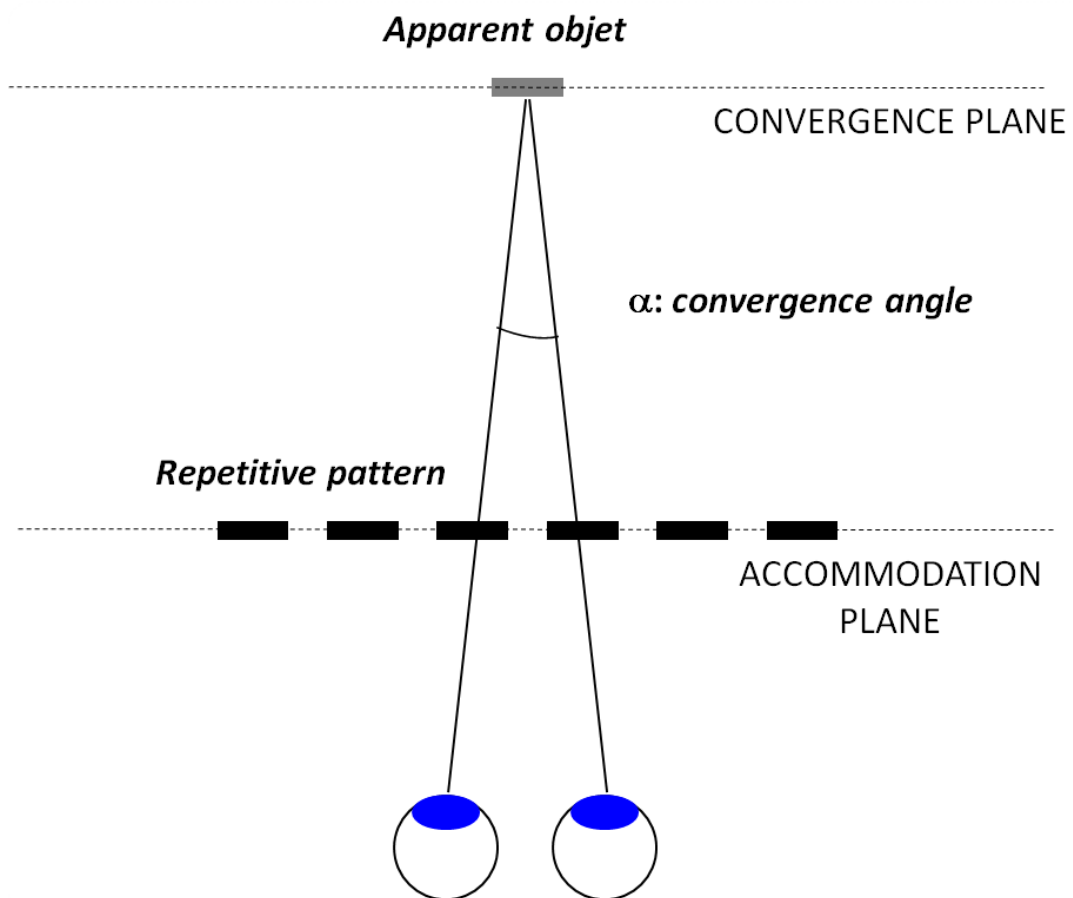
Figure 2. (A): Single image random dot stereogram (SIRDS), in which a DOG-2D (Difference of two 2-dimensional Gaussians) function may be observed. (B): Luminance map needed to generate the corresponding disparity or depth map.
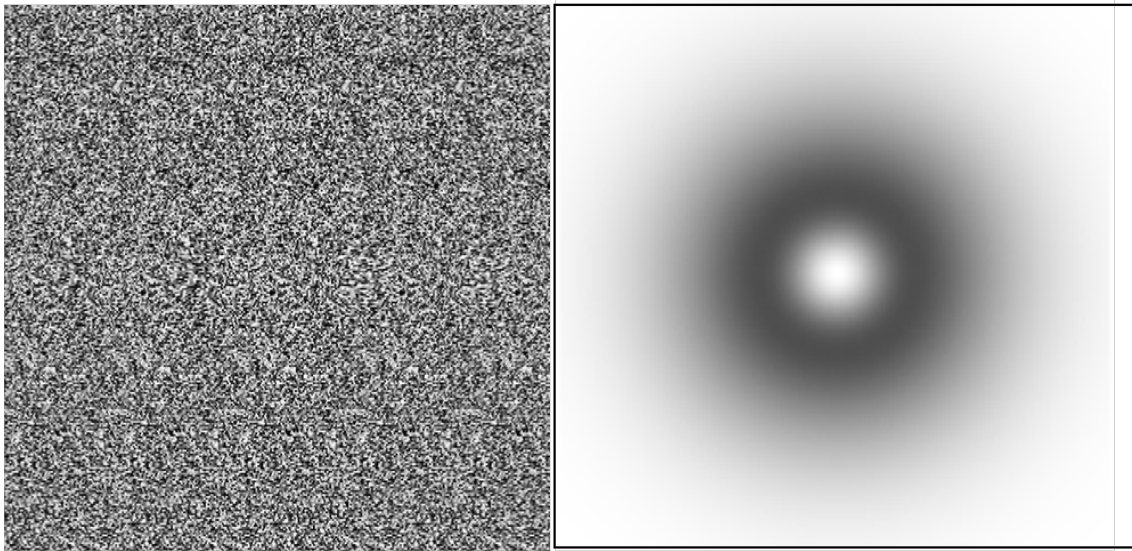
Figure 3. Minimum time required to perceive SIRDS. The median time of 10 seconds was selected as a cut-off value to categorize our participants.
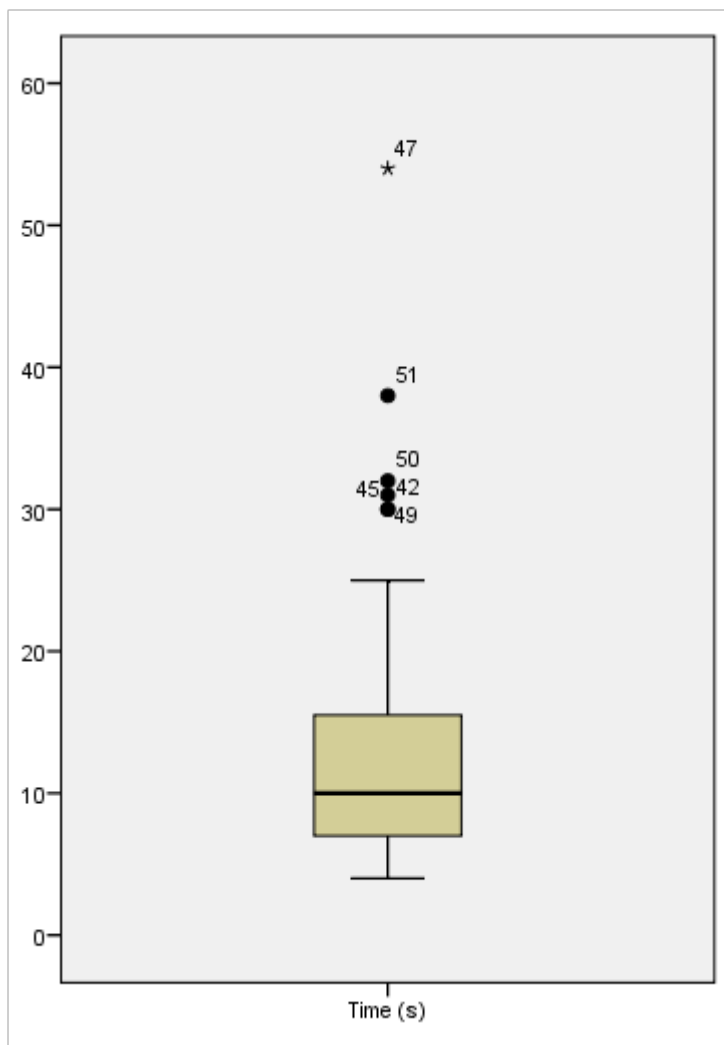
Figure 4. Location of each observer according to the values obtained from both canonical discriminant functions. The group centroid for each of the three skill level categories is also shown.