Psychophysics of Autostereogram Videos: Contrast, Repetition, Blur and Colour

Georgios Papadimitriou (papadingorge@gmail.com), Robert Fisher (rbf@inf.ed.ac.uk), Richard Shillcock (rcs@inf.ed.ac.uk)

School of Informatics, The University of Edinburgh, 10 Crichton Street, EH8 9AB, Edinburgh, United Kingdom

Abstract

Autostereograms are single image stereograms that take advantage of the binocular fusion and stereopsis of the human vision system and enable us to visualize three dimensional objects or scenes that are embedded in two-dimensional images. Autostereograms can be either static or videos that are created from animated depth masks of objects or scenes. In this paper, we present experimental measurements of the psychophysical thresholds for 3-D perception in Random Dot Autostereogram (RDA) videos with respect to contrast, repetition of the random dots that constitute the repetitive patches inside a RDA video, blur and colour of random dots. The approach we followed focused on performance data gathering for stereopsis achievement in human participants by conducting experiments on whether and how fast it is achieved. The stimuli used were autostereogram videos of basic objects (cubes, tubes, pyramids, disks, stars and pentagons) in which we varied the setting of one of the aforementioned features (contrast, repetition, blur and colour) while keeping the rest fixed. Our findings showed that there is an upper threshold of Gaussian blur radius at 33-35 pixels (≈ 2.38° – 2.53° visual angle), a lower threshold of 0.03 Michelson contrast, an optimal range of settings 70-100 pixels (corresponding to ≈ 5.05° – 7.22° visual angle) for repetition period and a lower threshold for luminance value of red/green random dots between 0.73 cd/m² to 0.98 cd/m² beyond which stereopsis is not achieved in human participants. The experiments, in addition to establishing the range of perceptibility, also demonstrate that stereo fusion is possible even when every frame in the video uses a different random dot field.

Keywords: Autostereogram Videos, Psychophysics, 3-D Perception, Blur, Contrast, Repetition Period, Colour

1. Introduction

Autostereograms are a type of stereograms that allow us to represent 3-D objects and scenes using ordinary display devices and means (computer screens, paper etc.) by taking advantage of the binocular fusion and stereopsis of the human vision system. They enable us to visualize 3-D objects inside a single 2-D image (see Figure 1) by fusion between repetitive random dots or textured patterns. Autostereogram videos (random dot or textured) are sequences of autostereograms, created by animated depth masks that represent objects or scenes. In our case, every frame in the video uses a different random dot pattern to render an underlying changing 3-D scene. Wheatstone (1838) discovered the stereoscopic vision and stereograms. The psychophysics of static and dynamic stereograms and most of their subcategories (Julesz, 1971; Julesz and Chang, 1976; Essig et al., 2004; Ee and Erkelens, 1996; Skrandies, 2009; Tanabe et al., 2005; Fujikado et al., 1998) is a well-researched field. On the other hand, there is no work known to us regarding the psychophysics and the space of perceptibility of autostereogram videos.

Although there are many studies on stereogram pairs (e.g., Julesz and Chang, 1976) there are few on autostereograms. Essig et al. (2004) studied vergence eye-movements in autostereogram images and the effect of image grain size (granularity) on such movements. They found that participants could not achieve stable 3-D perception of autostereogram images with large granularities. In addition to this, they found that regardless of the level of granularity in the autostereogram images, the participants performed divergence movements slower than convergence movements when they were trying to perceive 3-D objects inside the autostereogram images (Essig et al., 2004). Ee and Erkelens (1996) used stereograms to study the “temporal aspects of binocular slant perception in the presence and absence of a visual reference”. What they found was that when stereograms were observed for short time periods (less than few seconds) slant was poorly perceived in the case where no visual reference was present.

Skrandies (2009) assessed depth perception of adults with stereo-vision deficiency by using dynamic random dot stereogram pairs as stimuli by recording the neurons’ responses. What he found is hemifield differences between the right and the left parafoveal areas in information processing (i.e. “the processing of horizontally disparate stimuli information is more specific if it is processed in the right visual field than in the left visual field”). Tanabe et al. (2005) used dynamic random dot stereogram pairs to test the neural responses in Macaque brain areas that are related to vision. Their experiments showed that the stereoscopic depth representation in the V4 visual area is suited for “detecting fine structural features protruding from a background”. Fujikado et al. (1998) used dynamic random dot stereogram and coloured static stereogram pairs to assess stere-
2. Methods

We conducted the experiments at different time periods, using different numbers of subjects, different features/attributes of the autostereogram videos and slightly different pre-experimental data recorded for each participant.

2.1. Equipment and experimental setting

In both phases the stimuli (RDA videos) were projected with an Nvidia 9600GT M (default settings) on an 18-inch, 1800FP Dell LCD monitor of size 35.904×29 cm under the default settings. Another monitor (same model) was used for projecting an electronic chronometer which recorded the time each participant needed to identify the object inside each video. A chin support was used so that for each participant the head position and orientation would be the same throughout the experiment.

The experiments were conducted in a dim room with the only light sources to be the projection screen and the screen used for the electronic chronometer. The distance of the projection screen from each participant was fixed at 44 cm. Both the chin support and the projection screen were placed on the same plane. The height of the chin support was 36 cm and its width 6.4 cm. The projection screen formed an angle of 90° with the plane on which it was placed. The participant’s eyes formed an angle of ≈ 38° with a hypothetical plane that is perpendicular to the projection screen in its center.

2.2. Stimuli

Stimuli (autostereogram video) creation was a three step process. The first step created the 3-D objects used in the autostereogram videos. The second step created the animated depth masks from these objects and finally, the animated depth masks were used as an input for the autostereogram video creation. Six objects (disk, pyramid, star, cube, tube, pentagon) were created using 3-D Studio Max. Table 1 illustrates the creation parameters of these objects. The animated depth masks were created with 3-D Monster. Each animated mask consisted of 2000 frames, had a frame rate of 50 fps and a resolution of 640×480 pixels. The autostereogram videos created with the 3-D Miracle platform are divided into four categories: videos of different Gaussian blur radii, Michelson contrasts, random dot repetitions and colour of random dots (red/green). Michelson contrast is calculated by formula 1 (Arend, 2010).

\[ C_M = \frac{L_{\text{max}} - L_{\text{min}}}{L_{\text{max}} + L_{\text{min}}} \]

where, \( L_{\text{max}} \) is the constant maximum luminance (in our case the luminance of white dots ≈ 134.91 cd/m²), \( L_{\text{min}} \) is the minimum luminance (in our case the luminance of black dots varying from ≈ 6.86 – 125.88 cd/m²). Based on Table 3, formula 1 and the maximum luminance given above, the minimum luminance can be easily calculated in each case.

For each set of experiments we varied the parameter in question while keeping the rest fixed. Table 2 summarises the fixed generation parameters of the different Gaussian blur RDA videos. The values shown in Table 2 were determined experimentally with the aim of finding parameters that give good quality videos. The fixed parameters for the three remaining categories of autostereogram videos (Michelson contrast, repetition and colour) are essentially the same as the ones shown in Table 2 except the values of the parameter tested each time. All the videos used in our experiments were created using black &
white dots apart from the ones that belong to the colour category where the dots used are red & green. The values of each varying feature/parameter in each category of autostereogram video is shown in Table 3. Sample videos of the ones that were used during the experiments can be found at website1 while depth mask snapshots of the six test objects/scenes are shown in Figure 2.

Apart from the main stimuli for the experiments that are presented above, a blank video (no object was present) was used for each category of videos (five blank videos in total). These videos were used as ‘truth test’ videos due to their nature in order to further ensure that the participants were not making wild guesses regarding the objects shown in the videos. The fixed creation parameters of these videos were the same as the ones used for the main stimuli of each category.

2.3. Experimental Procedure

Our experiments are divided into two phases. The experimental procedures followed in each phase are similar to each other but distinct as described below.

2.3.1. Phase I

The performance of twenty six paid participants was recorded and analysed in the phase I experiments. The pre-experimental data recorded for each participant was their sex (16 males, 10 females), their handedness (22 right-handed, 4 left-handed), their age (19-57), their eye condition with respect to if the participant was wearing glasses/contact lenses (11 wore glasses/contact lenses and had myopia), any other known eye problems (1 was red/green color blind) and the experience of stereograms. Participants that failed the test were allowed to participate in the experiments but their results were excluded from the analysis. 26 people effectively took part in the experiments out of the initial 28.

The next step was the projection of autostereogram videos to each participant. The experiments of phase I can be divided into two sub-experiments. (i) different Michelson contrasts were used and (ii) different dot patch repetitions were projected. These two sub-experiments were conducted in the order mentioned above and each lasted approximately 20 minutes without intermissions in-between. For each corresponding values shown in Table 3 two videos of different objects were projected to each participant until the whole range was covered (e.g. for 0.03 Michelson contrast two videos were projected to each participant, one illustrating a disk and one illustrating a pyramid). The sequence of the videos with which each participant was presented was fixed but unsystematic.

The videos were projected in full screen (640x480 resolution with 1 dot per pixel) and before each video each participant was presented with an eight-second video of random black and white dots apart from the ones that belong to the colour category where the dots used are red & green. The values of each varying feature/parameter in each category of autostereogram video is shown in Table 3. Sample videos of the ones that were used during the experiments can be found at website1 while depth mask snapshots of the six test objects/scenes are shown in Figure 2.

Apart from the main stimuli for the experiments that are presented above, a blank video (no object was present) was used for each category of videos (five blank videos in total). These videos were used as ‘truth test’ videos due to their nature in order to further ensure that the participants were not making wild guesses regarding the objects shown in the videos. The fixed creation parameters of these videos were the same as the ones used for the main stimuli of each category.

2.3. Experimental Procedure

Our experiments are divided into two phases. The experimental procedures followed in each phase are similar to each other but distinct as described below.

---

1 http://homepages.inf.ed.ac.uk/rbf/SIRDVIDEOS/
white dot noise created in Matlab with 640×480 resolution and a two-second “get ready” message with a white font on a black background. This was done to help the participants lose focus from the previous autostereogram projection and prepare them for the next one respectively. For each autostereogram video the information recorded was: the object perceived by the participant and whether it was identified correctly or not, the time to achieve perception and whether the perception was stable or not in terms of whether the participant lost perception of the object for some time (unstable perception) or not (stable perception). The time to achieve perception (stereopsis) was measured by an electronic stopwatch which the participant stopped each time he/she perceived an object by clicking on the stop button, without looking at the stop watch. Each time the participant stopped the stopwatch we stopped the video projection to record time and what he/she saw. The perception achieved was divided into three categories: “no perception” when the participant had no perception of any kind of the five objects that were used for the experiments (cube, disk, pentagon, tube, pyramid), “something moving” when the participant stated that they could see a 3-D object but could not resolve what the object was and “object name” when the participant was able to identify what the object was (i.e. disk, cube etc.). A participant was considered to be able to identify an object under a specific setting for one of the two features (Michelson contrast, repetition period) if he/she was able to identify the object in one of the two videos of the same setting projected to him. Before proceeding to the next video the participant was needed to clear the stopwatch measurement by clicking the clear button. This was an additional way for the participant to lose focus after 3-D perception was achieved and then look back to the screen in which videos were projected.

### 2.3.2. Phase II

The performance of thirteen unpaid subjects was recorded and analysed part in the phase II experiments. The pre-experimental data recorded for each participant was the same as that in phase I. Additionally, we also recorded the score they achieved when performing a TNO stereo test prior to the experiment (stereo-acuity scores were in the range of 60-30 seconds of arc). One of the participants, despite not having any eye problems, viewed the autostereogram videos inverted (inside-out). Consequently this participant was excluded from the analysis of the experimental data.

Prior to the projection of autostereogram videos participants were tested for achievement of stereopsis using random dot autostereogram videos, similar to phase I. The participants who failed in the TNO stereo test and/or failed to recognize the objects inside the RDA videos were excluded from the experiments. As a result the performance of thirteen participants of the total fifteen volunteers was used for analysis.

The rest of the procedure remained relatively the same as in phase I with some differences. Phase II experiments are divided into two sub-experiments: (i) we used different amounts of Gaussian Blur and (ii) we used equiluminant (as measured by spectrometer) coloured dots (red/green) instead of black and white dots that were used in every other sub-experiment of both phases. Gaussian blur was generated using convolution with a mask whose standard deviation parameter was as specified. After convolution, the contrast of the video was renormalised to be the same as the un-blurred video. The sub-experiments were conducted in the sequence mentioned above and each one lasted approximately twenty minutes with no intermissions in-between.

### 3. Results

In this section we present the results of both phases. These results were analysed after removing the statistical outliers from the data obtained from the experiments. The results are presented as plots showing the percentage of videos for which stereopsis was achieved by the participants for different values of the varying parameter. Some of the plots are presented on a logarithm scale (fig. 3 & fig. 6) to observe a better trend.

#### 3.1. Phase I Contrast Results

Figure 3 shows the percentage of the participants who achieved stereopsis at different values of the Michelson contrast. The most important observation from figure 3 is that no participant was able to perceive the objects in videos of 0.03 Michelson contrast (-3.5066 on log scale). This finding shows us that people are not able to perceive RDA videos below a threshold that lies within the 0.03-0.06 range of Michelson contrast. For higher values of Michelson contrast, the performance of the participants is stable. It is also observed that the rate of improvement in the performance of the participants is higher for contrasts within the range of 0.06-0.12.

A participant is considered to have identified videos of a specific contrast setting either when he/she identified the object in one or both videos of the same setting. For videos of higher contrasts, the percentage of videos for which stereopsis was achieved by the participants is higher than 90% in each case.
3.2. Phase I Repetition Results

Figure 4 shows the percentage of videos for which stereopsis was achieved for different dot patch repetitions.

From Figure 4 it is observed that the optimal dot patch repetition for the RDA videos tested is in the range 70-100 pixels. Outside this range, the performance of the participants is found to decrease. In addition, it is observed from the statistics that the time for a participant to identify the object inside the video is initially high but decreases as the dot patch repetition increases until a repetition of 70 pixels is reached ($\approx 4.86^\circ$ visual angle). For repetitions greater than 70 pixels, the time for a participant to identify the objects stabilizes until a repetition of 100 pixels is reached ($\approx 6.94^\circ$ visual angle) and then, for greater repetitions, there is an increasing trend in time. These findings show us that the optimal dot patch repetition for the RDA videos tested is between 70 and 100 pixels ($\approx 4.86^\circ - 6.94^\circ$ visual angle). It is worth mentioning at this point that for repetitions of 60 pixels ($\approx 4.16^\circ$ visual angle) and below the participants stated that they observed the objects in a sliced form, (i.e. discontinuous reconstructions arising from incorrect correspondences).

3.3. Phase II Blur Results

The percentage of videos for which stereopsis was achieved and the values of blur radius is shown in figure 5. The most important observation we can make by examining Figure 5 is that no participant was able to perceive the objects in the autostereogram videos of 38 pixels blur radius (this corresponds to $\approx 2.64^\circ$ visual angle between the center of a hypothetical line that connects the participants’ eyes and the ends of the blur patch). Consequently, this enables us to define an upper threshold within the range of 37-38 pixels blur radius ($\approx 2.57^\circ - 2.64^\circ$ visual angle) above which human 3-D perception is not achievable with respect to RDA videos. It is also observed that the performance of the participants is more stable for lower values of blur radii and this stability decreased as the value of blur radius increased. This is as expected since the blurrier an object is, the harder for someone to identify it, let alone when this object is part of a RDA video where the observer has to perform binocular fusion to achieve stereopsis. The percentage of videos that were identified dropped drastically for blur radii greater than 34 pixels.

3.4. Phase II Colour Results

Figure 6 shows the percentage of videos for which stereopsis was achieved for different luminance of red/green random dots with the values of luminance on a logarithmic scale. The most important observation one can make by examining Figure 6 is that there is a decreasing difficulty in identifying objects in RDA videos as the luminance of red and green dots increases. Figure 6 also shows that even slight differences in luminance of red/green dots affects perception and consequently
identification greatly since for a luminance of 0.98 cd/m² (~0.0202 on logarithmic scale) everyone was able to identify the objects under this setting while for luminance of 0.69, 0.7, 0.73 cd/m² the percentages of videos being identified dropped drastically (7.69%, 30.77%, 42.31%). This places the threshold value of luminance for red/green random dots to achieve stereopsis in human participants in the range 0.73 cd/m² to 0.98 cd/m².

4. Discussion

In this paper we presented experimental observations on the psychophysical aspects of autostereogram videos with respect to blur, contrast, repetition period and colour of the random dots.

The findings for autostereogram videos of different uniform blur denote that humans are unable to perceive them for a Gaussian blur radius greater than approximately 33-35 pixels (= 2.38° – 2.53° visual angle). This consequently defines the upper threshold in human stereo vision with respect to blur. In addition, as blur increases, there is an increasing time for the participants to identify the objects in the videos and a decreasing number of videos in which objects are identified.

The results for experiments conducted with videos of different Michelson contrasts show that people have a threshold around 0.03 Michelson contrast below which they are unable to perceive the objects inside the autostereogram videos. There is a decrease in the time needed by the participants to identify the objects for contrasts up to 0.08. For higher contrasts the participants’ performance (the time to identification) is relatively stable.

Based on the findings for autostereogram videos of different repetition periods, there seems to be an optimal range of repetition period settings (70-100 pixels corresponding to 5.05° – 7.22° visual angle) outside which the performance of the observers becomes worse with respect to the time they need to identify the objects inside the videos. This leads us to conclude that the optimal parameters are the ones found in the middle of the parameter range (i.e. neither high nor low repetition periods are optimal for perception). In addition, there is a range of repetition period settings (30-100 pixels corresponding to 2.16° – 7.22° visual angle) outside which the observers either start misidentifying the objects in the videos or not perceiving them at all.

The experiments conducted with different luminance values of red/green colours of random dots provided with the lower bound for the luminance value in the range 0.73 cd/m² to 0.98 cd/m² to achieve stereopsis. It is also observed that above this range, as the value of luminance increases, more stable stereopsis is achieved in less time. The threshold below which stereopsis is not achieved is at 0.73 cd/m².

In addition to the above analysis, we also compared the performance between different groups of participants i.e. with respect to the sex of the participants, participants that wore glasses or contact lenses and participants that did not and finally between participants that were experienced in viewing stereograms (any type) and participants that were not. No statistically significant difference was observed in performance in the above groups except the case of experienced participants when compared to in-experienced ones. The t-tests performed to the recorded times of both groups resulted in small but statistically significant differences which show that experienced observers are significantly faster in identifying the objects and achieving stereopsis.

Although there has been much previous research on dynamic binary stereograms and some using autostereograms, we are not aware of research combining both, namely autostereogram videos. Here, we have observed perceptual thresholds for blur, contrast, colour contrast and repetition rate for autostereograms. However, irrespective of the perceptual thresholds, what is most surprising is the stability of the 3D percept even when every frame (presented at 50 fps) has a completely different random dot field.

References


