

Analyzing autostereoscopic environment configurations for the design of videogames

José Martínez Sotoca  · Miguel Chover  ·
Inmaculada Remolar  · Ricardo Loreto

Received: date / Accepted: date

Abstract Stereoscopic devices are becoming more popular every day. The 3D visualization that these displays offer is being used by videogame designers in order to enhance the user's game experience. Autostereoscopic monitors offer the possibility of obtaining this 3D visualization without the need for extra device. This fact makes them more attractive to videogame developers. However, the configuration of the cameras that make it possible to obtain an immersive 3D visualization inside the game is still an open problem. In this paper, some system configurations that create autostereoscopic visualization in a 3D game engine were evaluated in order to obtain a good accommodation of the user experience with the game. To achieve this, user tests that take into account the movement of the player were carried out to evaluate different camera configurations, namely, dynamic and static converging optical axis and parallel optical axis. The purpose of these tests is to evaluate the user experience regarding visual discomfort resulting from the movement of the objects, with the purpose of assessing the preference for one configuration or the other. The results show that the users tend to have a preference trend for the parallel optical axis configuration set. This configuration seems to be optimal because the area where the moving objects are focused is deeper than in the other configurations.

Keywords Autostereoscopic displays; Three-dimensional displays; videogames; multi-view display

1 Introduction

Autostereoscopic monitors are one of the most popular emerging technologies for use in visualization devices. This is due to the fact that these displays provide a 3D image without the viewer needing to wear any special viewing device. The consoles produced by Nintendo (3DS, new 3DS, etc.) have managed to bring autostereoscopic models to the market, which is an indication that these technologies are being considered as a novel means to create and visualize contents especially in the field of videogames.

Multi-view autostereoscopic televisions (MVATV) allow different people to simultaneously see the same content in 3D from different views of the scene in real time without any additional device, unlike other stereoscopic technologies that require multi-player connection and 3D glasses. However, several problems appear in the generation of content to be displayed on these MVATV, especially when a game is played using these displays. Several studies have been conducted that address the visual discomfort produced by displaying the movement of objects [1–3].

The comfortable viewing zone, which determines the threshold of the distance between the virtual object and the screen that would not induce visual discomfort, has thus been investigated and defined in different papers [1, 2, 5], most of which have focused on stereoscopic still images. However, other studies demonstrate that fast motion can also induce visual discomfort even if the object is within this comfortable viewing zone [2–4].

Besides this visual inconvenience, some issues that affect the user experience have to be taken into account when performing an autostereoscopic display, such as eye strain [6], crosstalk or ghosting [7], simulator sickness [8], etc. These symptoms make the game uncomfortable for users, and so they have to be analyzed and their possible effects avoided. This study focuses on finding an optimal configuration of the cameras that makes autostereoscopic visualization possible, taking into account the usual movement that occurs in most videogames.

This study considers several methods of configuring the cameras in a game engine to improve the user experience. In order to analyze the quality of the display with respect to the player's movement, a third-person action-based videogame was designed. The type of videogame chosen was an Autostereoscopic Shooter (AS) game. The aim is to analyze the camera setup that best reduces the problems of accommodation-convergence that are generated in the users when there are changes in the movement of the player. Furthermore, an evaluation questionnaire is designed in order to collect the user's opinions when evaluating the quality of the visualization of the game. Results are evaluated and an optimal camera configuration is proposed.

The rest of the paper is organized as follows. Section 2 provides a concise overview of the state of the art related to auto-stereoscopy, videogames and user experiences with this type of displays. This section briefly analyzes the guidelines for improving the quality of the user experience in viewing via stereoscopic media, and establishes a hypothesis to be tested in this article. Section 3 describes the design of the different configurations that are going to be tested. This section also describes the game and the building of the interlaced image to be displayed on a 3D multi-view Autostereoscopic Television. Section 4 explains the tests performed and analyzes the results by means of a statistical analysis for the three camera configuration methods. Finally, the conclusions and some future work are discussed in section 5.

2 State of the art

The research on the relationship between the user and auto-stereoscopy involves several areas that have to be taken into account. They cover the study of problems regarding the physiological and physical effects on human beings that involve the quality of the user experience. In addition, another field of research that needs to be addressed is how stereoscopy influences human capabilities and performance.

2.1 Convergence-accommodation conflict

The convergence-accommodation conflict is a widely documented problem caused by the eyes misleading the brain by forcing a change in the usual way of focusing the eye's lens, which gives rise to adverse effects. On the one hand, eye accommodation is the process by which the lenses are adapted to focus on an object which is carried out by means of the ciliary muscle. The eyes change the optical power by modifying the curvature of the lens and making the object look focused [6] [11]. On the other hand, convergence is the process by which both eyes move in opposite directions, so that the object to be focused on is adjusted on the retina [10]. The object is said to be in focus when the visual axes are aligned at that point.

When a person looks at a point, the accommodation and convergence distances are the same. This also happens with the stimulus that goes to the accommodation and convergence systems [12]: the eyes focus on and converge to the plane where this stimulus is. When a human is looking at an autostereoscopic monitor, the eyes are focused on the monitor and the distance to it is constant [4].

The brain then receives *planarity* information and absence of 3D. Nevertheless, each eye receives a different image that produces the 3D effect. In this case, the convergence distance is different from the focus distance, as shown in Figure 1. Some research demonstrates that constantly favoring the convergence-accommodation conflict does not cause visual fatigue if the difference between the distances is maintained within a certain range [14] [9].

2.2 Effects of virtual movement on the display of stereoscopic contents

There are two main types of movement in stereoscopic displays: movement on the plane and movement in depth. Planar motion means that the object moves only on a plane parallel to the screen. In-depth motion is defined as that which occurs when the object moves toward or away from the observer. Some research has determined that planar movement combined with in-depth movement of objects causes visual fatigue [13, 15].

When displaying 3D images on flat stereoscopic screens, the binocular angular disparity φ can be expressed by the degree of visual angle [13], as illustrated in Figure 2. Let F be the situation of the object on the screen and let α be the angle formed by the focus direction of each of the eyes when they focus on the screen. In this situation, when the objects are on the screen plane, the binocular angular disparity is 0 degrees which means there is no disparity between the two retinal images.

The movement in depth, as well as the planar movement, of static images also causes visual fatigue. Particularly, quick movements cause an increase in discomfort in dynamic images [4]. When dealing with the speed of objects in the scene, it is important to take into account changes in the binocular disparity. The binocular disparities φ_A and φ_B produced when the fixation point F moves to points A and B can be calculated by Eq. 1. Furthermore, if a frenzied scene is going to be viewed, it is recommended to have a small disparity or, what is the same, to maintain a small angle of difference between the images.

$$\varphi_A = \beta - \alpha \quad (1)$$

$$\varphi_B = \gamma - \alpha$$

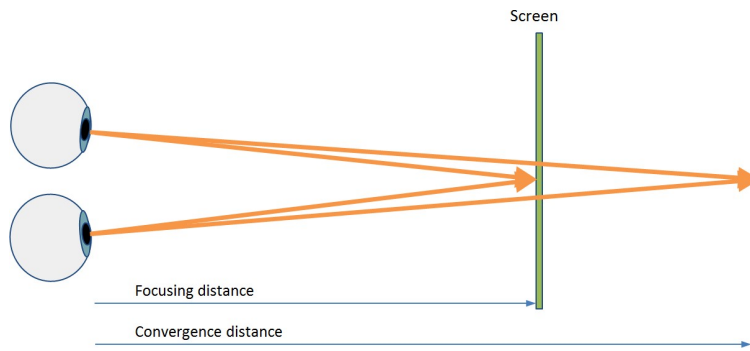


Fig. 1 Convergence-accommodation conflict in three-dimensional visualization.

Another concept that has to be taken into account is the distance between the cameras that are generating the images. In order to feel a realistic three-dimensional sensation, the viewer should receive different, slightly separated images for each eye obtained from a camera array. For that, it is necessary to consider the pupillary distance (PD) (Figure 2). The average adult pupillary distance is between 54 and 74 millimeters, while the PD of children is between 43 and 54 millimeters. However, if a videogame is going to be visualized, the best experience is sought, so this distance may vary for the sake of clarity and reduction of asthenopia. In this work a value of 65 millimeters was used to obtain the binocular disparities φ_A and φ_B .

2.3 Study of user experience viewing a 3D television

User eXperience (UX) is a difficult term to define in the sphere of human-computer interaction. The consensus states that UX is defined by three elements: the user, the system, and the context [17]. The user manipulates the system, which is the software required for the product to work, while the context represents the circumstances in which the activity takes place.

In our work, the target users are young people, who lose interest as the discomfort or physiological effects produced by 3D televisions (3DTV) increase. The contents that they demand are those typically required of a normal television, but they look for more than just objects contained in a box: they are interested in interactive contents and sharing their experience with other people. According to [18, 19], stereo monitors are causing a greater sensation of presence, immersion and depth in the visual contents.

Some studies [6,10,17] have investigated into the human capacity to deal with different tasks and the difference between solving them using a stereo system or from a single camera, concluding that sixty percent of the subjects solve the tasks better in the stereo field. However, in [20] it is stated that just a right amount of stereo is an important factor for a good quality experience, as it increases the sense of presence among users.

The literature allows the user to follow a series of guidelines for improving the quality of user experience in viewing stereoscopic media. The following are some of those listed:

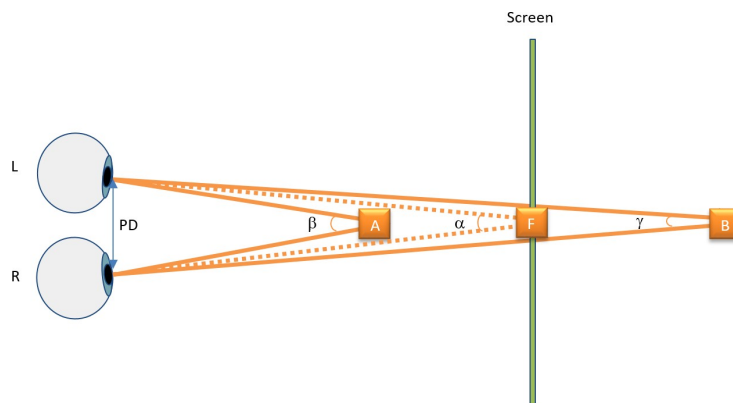


Fig. 2 Binocular angular disparity, where F is a point on the screen plane that moves to A and B .

1. The in-depth movements must be smooth and slow. The main action should be carried out in the comfortable viewing zone. If the camera array is focusing inside this zone, the in-depth movement that happens at a distance from this area is not important and can be frenetic. By following these guidelines, the convergence-accommodation conflict is potentially reduced [4,19,21,15].
2. It is important to maintain a correct separation between cameras and to obtain a render of the best possible quality, since the occurrence of visual artifacts in them leads to an increase in visual fatigue [21,22].
3. Depth in the videogame is given by elements in the medium and distant plane that have both depth and height [21, 23]. If the objects are situated too close to the camera, the movement can be frenetic, and if they are far away from it, the objects can appear blurred.
4. The viewing periods should be short because the human brain is not used to changing the focus on objects that move outside the comfort zone, so this can produce visual fatigue [16].
5. To create a dynamic and realistic effect of movement inside the comfortable zone, the game needs to know what the player is looking at [29].

Focusing the work on point 1, our hypothesis states that: *if there is a scene where the main action is kept within the comfort zone, leaving the plane of convergence static will favor the three-dimensional sensation and will provide the best user experience.*

3 Dynamic multi-view content generation

In this section, a study of the conditions of visualization and integration of synthetic scenarios in an autostereoscopic device is analyzed. In order to perform a depth analysis, these scenarios were rendered from a set of cameras with different static and dynamic configurations.

Firstly, a videogame was developed using two different scenarios. In both of them, all the action is performed in the center of the projected scene in the search for a greater sense of depth. The avatar that represents the player is a spacecraft. In the first scenario, the player moves inside a tube-shaped enclosure. At the same time, a set of asteroids appear in the videogame, emerging from the background and moving inside the tube-shaped enclosure so that the sensation of perspective is increased. Figure 3 shows a screen-shot of this scenario on the left. The other scenario designed for the second version of the game was a procedurally generated terrain. The spacecraft flies over the terrain trying to avoid hitting the asteroids that also emerge from the background. This scenario is shown on the right of Figure 3.

Initially, the activity of the player in the game is passive, that is, he/she only observes the game without interacting with it. The spacecraft moves with a non-constant acceleration approaching or moving away from the viewer. The player does not control the path his or her avatar takes from the start to the end (although he/she may be able to pause that movement). In this part, the player has to concentrate on the problems linked to the convergence-accommodation conflict of the 3DTV viewer regardless of the actions that the he/she can perform in the game. After this visualization period, the player interacts with the game, moving the spacecraft to avoid the asteroids that are going to crash into it. This part attempts to test the players' response in autostereoscopic visualization.

Several camera configurations that cause different multi-view results of the scene were tested. The values of the different parameters were obtained by analyzing the technical description of the 3D MVATV that is going to be used in the experiment (8-view 46-in xyZ 3D LCD-display with lenticular lenses). The final configuration is detailed in Section 3.1. Next, for each frame, the RGB information pattern was interlaced to be sent to the auto-stereoscopic display (Section 3.2). The generation of interlaced images was performed maintaining an interactive frame rate using GPU shaders. These last two steps were integrated inside a powerful game engine (Unity 3D) in order to allow interactive real-time visualization. Finally, the integration of information in Unity and the characteristics of the autostereoscopic display are detailed in Section 3.3.

3.1 Virtual world configuration and visualization

Camera setup is a crucial point for multi-view content generation. Therefore, depending on the configuration of the parameters in the camera model, different effects can be produced in the visualization of scenes. This fact makes it necessary to test different static and dynamic configurations that can be analyzed to improve the user experience in the game.

3.1.1 Configuration of the videogame

The game designed for this study was developed using a game engine in which coordinates system are measured in Graphics Units (GU). Moreover, the time unit is the time elapsed between two consecutive frames, *deltaTime* (dT), so, the velocity is given in GU/dT. The videogame runs at 30 frames per second and thus it can be established that $dT = 0.033$ seconds.

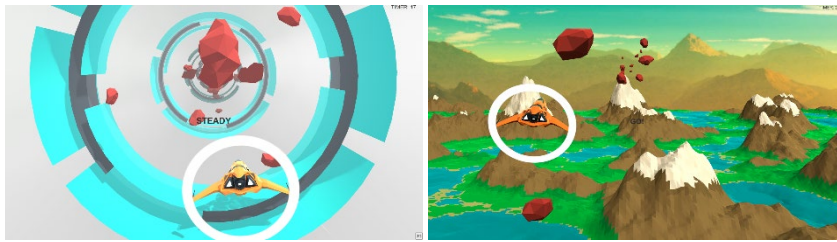


Fig. 3 Screenshots of the Autostereoscopic Shooter (AS) game developed with two different scenarios.

In order to experiment with the different camera configurations, some distances were set, as shown in Figure 4. First of all, the array of cameras was placed at 2.5 GU with respect to a plane where the objects are focused, situated on the television screen. The player was represented by a spacecraft that can approach or move away from the set of cameras within a range of depths. In the game, this movement was established from 1.3 to 8.5 GU taking the situation of the cameras as the initial point.

Another interval that has to be set is the user comfortable viewing zone, that is, the range of depths outside and inside the display where the objects can be seen with sufficient comfort by the viewer. This zone was established experimentally by visualizing a 0.5 GU cube with a checkered texture. Next, we gathered the opinion of some users located approximately 4 meters from the screen. This distance is in the range between 3 to 7 meters recommended by the manufacturer of the 3D MVATV. During the experiment, the cube was moved in-depth by locating it at different levels of depth, and the users determined when visual defocus is produced. The experiment established a comfortable viewing zone of between 1.04 and 3.45 GU.

The speed of the spacecraft can vary in this comfort zone, with 0.11 GU/dT for slow velocity and 0.23 GU/dT for normal velocity. At the same time, a set of asteroids appear in the videogame moving inside a tube-shaped enclosure or flying over a terrain, so that the sensation of perspective is increased. Their emission point was located at 100 GU, this distance being considered as the background. However, the viewers can recognize them when they reach a distance of approximately 20 GU from to the cameras.

To complete the configuration and make it easier to measure the visual discomfort, graphic units were converted into real measures. The scale factor to convert GU into meters was calculated experimentally using the checkered texture cube again. The fact that the size of the cube viewed by the user diminishes with the distance because of the perspective projection, thus maintaining a constant angle as taken into account. Accordingly, the size of the cube was measured when it represented on the screen, and when it moved off-screen at different depths. Considering the triangles formed by the vanishing point of the perspective projection and the measured height of the cube in these distances, by similarity of triangles, the Thales theorem was applied to establish a scale factor that relates in-depth measurements of the real world to the graphic world. This scale is about 0.31 meters for each GU.

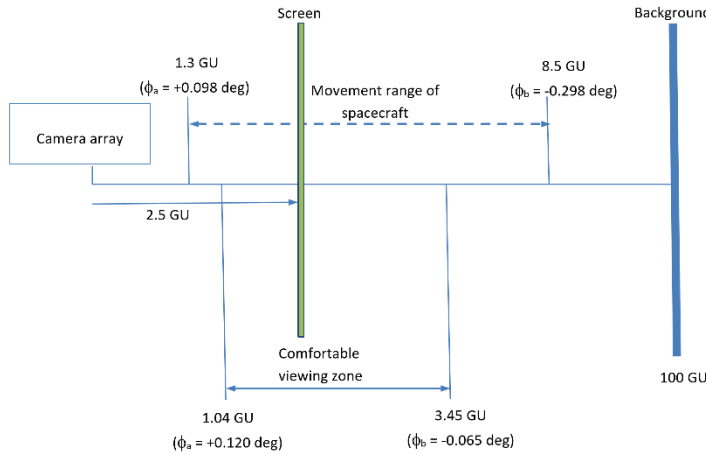


Fig. 4 Scheme of the videogame designed for performing the testing in graphic units and the binocular disparities in degrees.

In addition, it is possible to convert the intervals from GU to meters. The interval where the spacecraft moves in-depth was therefore updated to real measures, resulting in the range of [3.62, 5.88] meters. The binocular disparities φ_A and φ_B between these distances were calculated according to Equation 1, the results being +0.098 degrees in the position closest to the viewer and 0.298 degrees in the furthest one (see Figure 4). The comfortable viewing zone was established in the range of [3.54, 4.30] meters, the binocular disparities varying in the range of [+0.120, -0.065] degrees. Finally, the background where the asteroids emerge from, is located 34.60 meters away from the viewer.

3.1.2 Configuration of the set of cameras

In [24] the authors performed user tests in a First-Person game applying different camera setups, and showed that the configuration that produces more natural and realistic sensations was the planar configuration. Following the results of [24], a planar camera setup is presented in this work. A rack of cameras was disposed, where n linearly arranged cameras are pointing toward the scene. Several parameters were set to configure these cameras. In the case of the distance between them, the bigger the distance between the cameras is, the more parallax is obtained. Thus, the proposed configuration was to position eight cameras (number of views of the used 3D MVATV), the distance between the optical centers of two consecutive cameras in the camera array being set to 0.01 GU.

The distance to the convergence plane will vary statically or dynamically, and the orientation of the optical axes will be modifiable by varying the optical axes of the cameras from parallel to convergent. The spacecraft performs in-depth movements in the third-person game that was designed, so its distance to the camera position can vary within a range of distances. Moreover, in this kind of games, the movement of the focus point of the set of cameras must vary more smoothly than the movement of the player, since sudden changes in the images displayed by the set of cameras generate visual fatigue. This allows us to test three configuration methods:

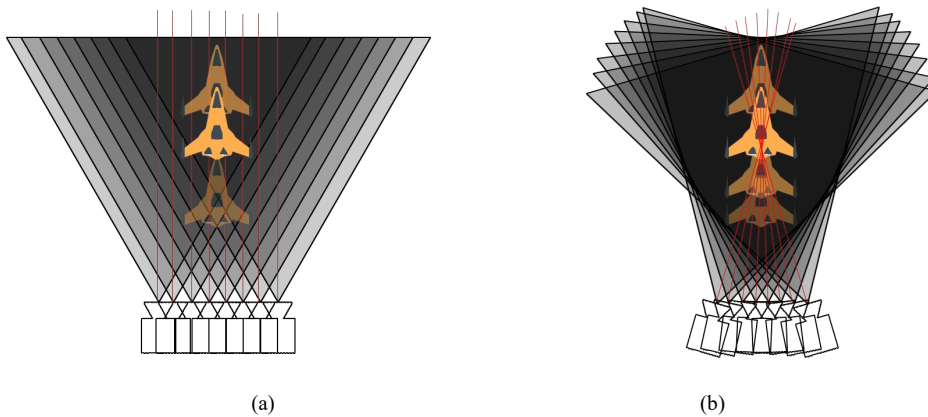


Fig. 5 (a) Parallel optical axes (POA). (b) Static converging optical axes (SCOA).

1. *Parallel optical axes (POA)*. In this configuration, the quality of the visualization will be determined by a horizontally distributed camera configuration, where the optical axes of the cameras are always parallel (Figure 5 (a)). This arrangement is the same planar setup as the one used in [24]. The Field of Vision (FOV) of the cameras was set to 60 degrees. In this configuration, as can be observed in the image, a common area appears in the scene at a certain distance from the cameras whose projected information can be seen in all of them. This area allows the convergence plane to be placed perpendicular to the optical axes of the cameras.
2. *Static converging optical axes (SCOA)*. This configuration was designed to evaluate the quality of the display that occurs when the optical axis is pointing to a static convergence point, situated on the display screen (Figure 5 (b)). This means that the distance from the set of cameras to the convergence plane remains at a constant value. In this case, the FOV was set to 41 degrees. This configuration is recommended by the manufacturer of the 3DTV.
3. *Dynamic converging optical axes (DCOA)*. In this configuration, the quality of the three-dimensional display will be evaluated with a planar camera arrangement where the optical axes are focusing on a point located in the plane of convergence. This point is indicated by the depth of the spacecraft relative to the cameras, so the convergence distance varies during viewing. In this experiment, the changes due to the velocity of objects were evaluated, taking into account the fact that the convergence plane varies its depth with the spacecraft's movement (see Figure 6). The FOV is also set to 41 degrees in this last configuration.

3.2 Generation of interlaced images

The 3D MVATV display is fitted with high precision lenticular lenses to generate the 3D sensation produced by different images being viewed by each eye. The 3D viewing is performed by combining the information obtained from the images taken by n -cameras. These images are initially stored in a single texture by applying a render-to-texture off-screen operation. They are then distributed in a new squared image composed of tiles.

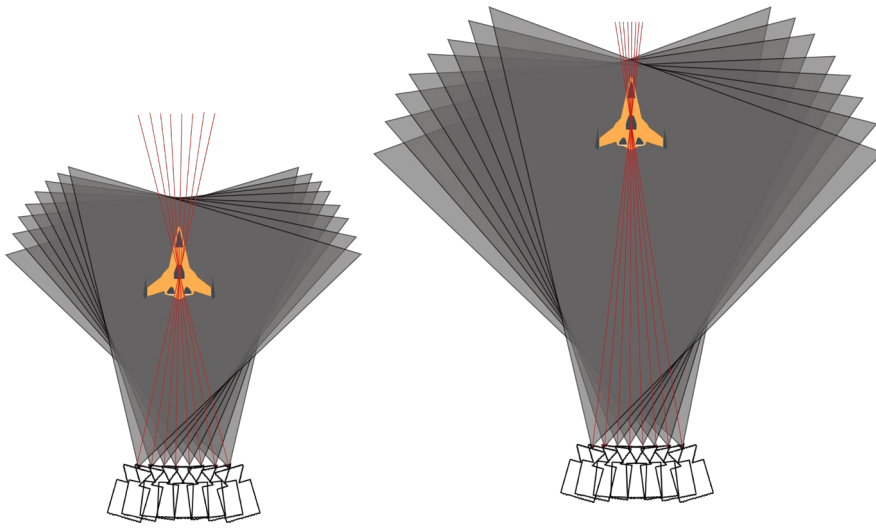


Fig. 6 Dynamic converging optical axes (DCOA). The convergence point varies depending on the situation of the player.

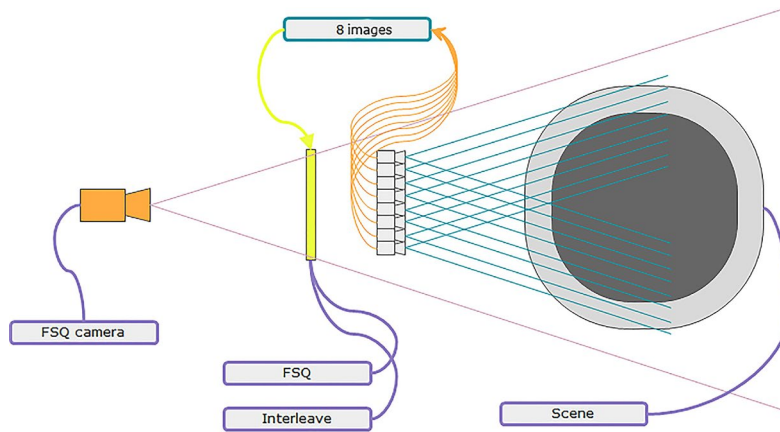


Fig. 7 Process of generating the final rendering in the 3D MVATV, using the POA configuration of the cameras as an example.

Lastly, this image has to be interlaced by distributing the tile information in each of the sub-pixel distributions in the 3D MVATV.

The composed image will also be stored in a texture and visualized in the 3DATV taking advantage of the graphics hardware. It will be applied to a quad through a shader, performing a technique called Full Screen Quad (FSQ) carried out in the vertex shader. This quad occupies the entire region of the viewport observed by a new camera. This whole process is described graphically in Figure 7.

The autostereoscopic device that was used in this work has eight views and its manufacturers recommend two possible configurations to compose the final texture with eight views: a mosaic of 3 3 tiles, where the information about one of the images has to be used twice, and another configuration of 4×2 images. Both configurations allow the information captured by the cameras to be interlaced. In our work, the arrangement of 4×2 images is the one on which the interlacing algorithm was developed, based on the properties of the display used in the visualization (Figure 8). The images obtained from each of the cameras have a resolution of 720×360, which, after performing the composition, resulted in a 1440×1440 square image.

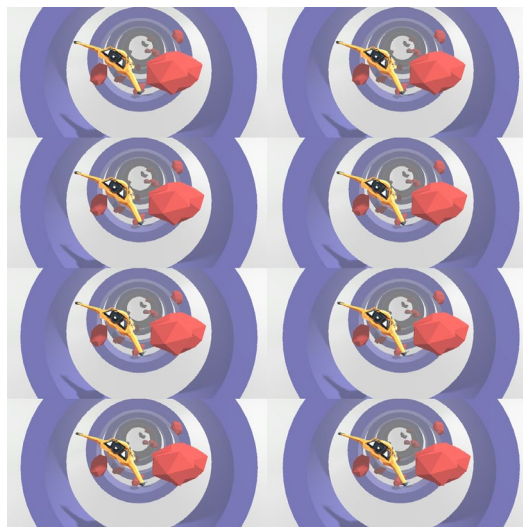


Fig. 8 Image of the square tiled texture (1440×1440) composed from the camera images, each with a resolution of 720×360.

R	G	B	R	G	B	R	G	B	R	G	B	
8	7	6	5	4	3	2	1	8	7	6	5	...
1	8	7	6	5	4	3	2	1	8	7	6	...
2	1	8	7	6	5	4	3	2	1	8	7	...
2	1	8	7	6	5	4	3	2	1	8	7	...
3	2	1	8	7	6	5	4	3	2	1	8	...
4	3	2	1	8	7	6	5	4	3	2	1	...
4	3	2	1	8	7	6	5	4	3	2	1	...
5	4	3	2	1	8	7	6	5	4	3	2	...
6	5	4	3	2	1	8	7	6	5	4	3	...
...

Fig. 9 Pattern to access the cameras and obtain the R, G or B values of the interlaced image.

The 3D MVATV used in our experiment has a sub-pixel distribution that combines the RGB values of different cameras. Figure 9 shows this sub-pixel distribution with the number of the camera that is employed in each of the three color channels, RGB. If we observe this pattern, it can be seen how a repetition of two rows is produced (rows 3 and 4, and rows 6 and 7 in our example). As an example of how the distribution is combined, Figure 10 (a) shows the tiled image resulting from composing the possible images obtained from each of the eight cameras. To simplify the explanation, different colors were applied to the image obtained from each camera, thereby ensuring that each color channel can take values from 0 to 255. Moreover, each camera is identified by a number in the tile. The result of interlacing these images is shown in Figure 10 (b).

As can be seen in this figure, the values of some rows are repeated. The method used by the 3D MVATV to compensate for this repetition is to interpolate all the repeated rows with neighboring row that is situated just above or below it. The resulting image is shown in Figure 11. It can now be seen that the transition of colors is smoother. Moreover, on observing the pattern of access to the image cameras (Figure 9), a displacement can be seen. Apart from the repeated rows, each row in the pattern access is displaced one position to the right in order to compose the next one.

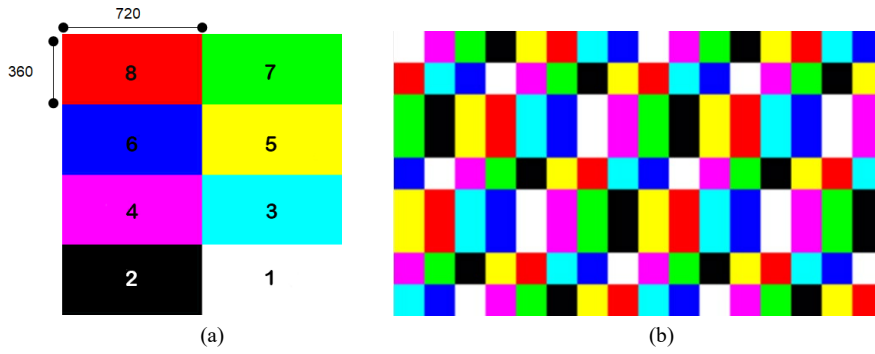


Fig. 10 (a) Tiled texture where each tile is represented by a color. (b) Interlaced image obtained following this pattern.

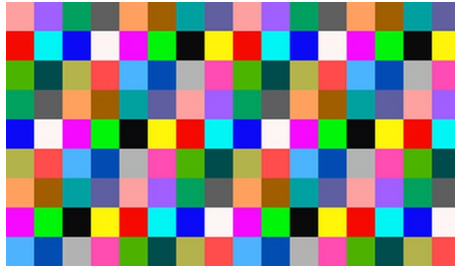


Fig. 11 Final image obtained after performing the interpolation of the required rows.

This arrangement is required due to the fact that the semi-cylindrical lenses are arranged in an inclined manner to reduce the Moiré effect following a slanted parallax barrier distribution [28]. The slanted parallax barrier covers the display and defines a particular light direction for each sub-pixel. Depending on the viewing angle and the distance of the observer from the display, most of the sub-pixels are masked. The structure of the optical filter defines a certain correspondence map between sub-pixels and views.

3.2.1 Implementation to obtain the interlaced image

The information included in Table 1 describes the variables that had to be defined to perform the interlacing. The square tiled texture is stored in the *Main_Tex* variable. In addition, as input parameters some more data are required: the dimensions of the square texture, *Tiled_dim*; the width and height of each image of the camera, *img_w* and *img_h*; the width and height of the FSQ rendered in the 3DATV, *w* and *h*, the interlaced image will be sent; and the number of patterns or cameras, *PatternLength*, engaged in the tiled texture. With this information, a set of functions are defined that will be used during the Fragment Shader process to interlace the content following the sub-pixel distribution.

The *GetPatternRGB* function (Algorithm 1) calculates the RGB values of the texture for a position determined by the row and the column input variables. Next, the displacement of the pattern is calculated so that the RGB values stop being sequential and they are adapted to those shown previously. Finally, to obtain the correct pattern, an adjustment between the pattern and the displacement is performed, which involves capturing the new RGB values and performing the interpolation of the rows that are required.

Table 1: Variables required by the code that was implemented.

Variables	Description
<i>Main_Tex</i>	tiled texture
<i>Tiled_dim</i>	texture dimensions of the tiled image
<i>img_w, img_h</i>	width and height of the camera images
<i>w, h</i>	width and height of the FSQ
<i>PatternLength</i>	number of patterns or cameras

The *GetCamera* function receives the tile number, a number in the interval $[1, 8]$, because the system requires eight cameras to perform the autostereoscopy. This function returns the position (row, column) of the image rendered from that camera in the non-interlaced texture. And finally, the *GetRGBValue* function (Algorithm 2) reads the color values of each image obtained from a camera in the tiled texture. All these functions are used in the main algorithm implemented to compose the interlaced image. This is shown in Algorithm 3.

Algorithm 1: Function *GetPatternRGB*. By giving a row and column of the tiled texture the new RGB values are estimated.

```

1 Function GetPatternRGB (int row, int column)
2   int normalPattern = (3 * ((w - row - 1) * h + column)) mod PatternLength;
3   int displacement = (2 * (row) / 3) mod PatternLength;
4   int adjustment = (normalPattern - displacement);
5   int newR = adjustment < 0 ? PatternLength + adjustment : adjustment;
6   int newG = newR + 1 > (PatternLength-1) ? 0 : newR + 1;
7   int newB = newG + 1 > (PatternLength-1) ? 0 : newG + 1;
8 return int3(newR, newG, newB)

```

Algorithm 2: Function *GetRGBValue*. This locates the required value on the tile and assigns it to the indicated value channel, 3: Red, 2: Green, 1: Blue.

```

1 Function GetRGBValue (int tile, int row_cam, int col_cam, int RGB)
2   col_cam = ceil((col_cam - RGB) / PatternLength);
3   pixelcam = float2(GetCamera(tile).y + col_cam,
4                     GetCamera(tile).x - row_cam - 1);
5   texelValue = tex2D(Main_Tex, float2(pixelcam.x, pixelcam.y) / Tiled_dim);
6 return (RGB==3 ? texelValue.r : RGB==2 ? texelValue.g : texelValue.b);

```

Algorithm 3 presents the pseudocode of the shader that generates the interlaced image from the tiled texture. The implementation was performed in the Unity 3D engine using the HLSL shading language.

Algorithm 3: Algorithm for obtaining the interlaced image.

```

1 Function frag(v2 f IN) : Color
2 // Obtain the number of the row and column of the texture analyzed
3   row = (1.0 - IN.uv.y) * h
4   column = round(IN.uv.x * w)
5 // Obtain the RGB value in the indicated row and column
6   Pattern RGB = GetPatternRGB(row, column);
7 // Calculate the appropriate row and column to perform the interpolation
8   row_Cam = (row)/3;
9   col_Cam = (column)*3.0;
10 // Calculate the RGB value in the final image
11  R = GetRGBValue(Pattern_RGB.r, row_Cam, col_Cam, 3.0);
12  G = GetRGBValue(Pattern_RGB.g, row_Cam, col_Cam, 2.0);
13  B = GetRGBValue(Pattern_RGB.b, row_Cam, col_Cam, 1.0),
14 Color = float4(R, G, B, 1);
15 return Color;

```

Figure 12 illustrates the interlaced image obtained by the shader. The input texture is the one shown in Figure 8.

3.3 Integration in Unity:

The videogame used in the test was developed with a powerful game engine: Unity 3D [26]. This allows us to use the capabilities of this multiplatform tool to integrate interactive content and apply GPU shaders. To conduct the experiments in the different POA, SCOA and DCOA configurations, the following tasks were performed:

1. Development of a script that, starting from a camera with pre-set values for the FOV and for the near and far clipping planes, can make copies in order to create a set of linearly disposed cameras. Depending on the selected camera configuration (POA, SCOA and DCOA), another script was developed that establishes the convergence plane. In this script, the parameters such as FOV, distance between cameras, and distance to the convergence plane can be modified by means of a keyboard.
2. Unity generates a quad that has an associated shader to perform the interlacing. It is linked to the viewport of the camera that visualizes the scene using the FSQ technique with that interlaced image. The distance of this camera with respect to the set of cameras always remains fixed.
3. The game control script changes after a certain period of time between the visualization of the game demo, where the player only observes the scene, and the interactive videogame, where he/she can operate the spacecraft and avoid the asteroids.

3.4 Characteristics of the display

The different setups were displayed in the 8-view 46-in xyZ 3D LCD display with lenticular lenses. The lenticular barrier is a transparent sheet composed of evenly spaced cylindrical lenses. The sub-pixel distribution of the 3DTV or diffusion screen, together with the lenticular barrier, allows certain lighting information to be visualized or hidden, thereby causing the 3D effect [25]. The diffusion screen follows a specific order to access the RGB information of the cameras, which is detailed in Figure 9. The result is an RGB sub-pixel pattern where the information can be found in a region of 3×8 colors of the final image.



Fig. 12 Interlaced image generated for display in the auto-stereoscopic device.

This type of monitor has an optimal viewing distance that ranges between 3 and 7 meters with respect to the television. In addition, the viewing angle is 100 degrees, which implies that if a user is at an angle of inclination above 50 degrees. Furthermore, this device has horizontal parallax and its resolution is 1920×1080 Full HD. With this horizontal parallax, when viewers move from side to side, objects farther distances appear to move more slowly than objects close to the viewers.

The information from the eight cameras was stored in a mosaic configuration of four rows and two columns, where the resolution of the image of each camera was 720×360 pixels, generating a square image of 1440×1440 pixels. The low resolution of the images used for interlacing can induce aliasing, which has a detrimental effect on the quality of the display. However, by applying an antialiasing processing to the 1440×1440-pixel image that contains the eight images of 720×360-pixel images, this is no longer appreciable.

4 Experiments and Results

Several tests were designed to confirm the initial hypothesis that assumes that if the plane of convergence does not vary, the 3D sensation is better. They were designed to measure the playability [30,31] and the correct visualization of the videogame [32]. These tests evaluate each of the experiences offered by the three configuration models in relation to these characteristics: playability of the game, annoyance conditioned by the visualization and finally, preference of the configuration (POA, SCOA and DCOA). To make it possible to reproduce the experiments and continue with this research, all the algorithms and the data involved in these tests can be download from [33]. Some instructions about managing the interactive environments have also been included there in order to know how to play the developed videogames.

A total of 45 students who considered themselves regular players, served as subjects for this experiment. Their ages ranged between 18 and 35 years. Concerning to the gender, 26 of them were men and 19 women. All of them had normal or corrected-to-normal visual acuity. Previously, they were informed about the aim and the method of the experiment, and they gave their informed consent to take part in this study. Regarding the level of previous experience of playing digital games, all of them stated that they had used videogames since a very young age. Finally, concerning the visualization of 3D videos, they all informed that they had poor experience in this field.

4.1 Experiment and questionnaire

In this experiment, each user had to guide a spacecraft to prevent it from crashing into the asteroids. They play the game, moving the avatar in all directions, in the two different scenarios that were designed. This process was repeated for each of the three camera configurations. At the end of the experiment, all of them were asked to fill out an evaluation questionnaire in order to express their sensations when faced with each configuration.

The questionnaire was designed by combining the questions presented by Shibata et al. [32] to measure visual discomfort with stereo devices and those put forward by Brockmyer et al. [30] and Nacke et al. [31] for evaluating playability. The final questionnaire is shown in Table 2. It contains questions about Playability of the game (P1...P6), Annoyance (A1...A5) and the preference of the Configuration (C1, C2). The response to each question was evaluated on a 5-point Likert scale. Regarding the three camera configurations, 5 is the score that indicates the best experience, and 1 indicates the worst.

Table 2 Questions presented to the test participants to evaluate the QoE.

	POA	SCOA	DCOA
Playability			
P1. Things seem to be unpredictable	—	—	—
P2. Playing seems automatic	—	—	—
P3. I lose track of where I am	—	—	—
P4. I play without thinking about how to play	—	—	—
P5. I really get into the game	—	—	—
P6. I feel like I just can't stop playing	—	—	—
Annoyance			
A1. How much seasickness do you feel?	—	—	—
A2. How tired and sore are your neck and back?	—	—	—
A3. How do your eyes feel after playing?	—	—	—
A4. Which session was most fatiguing?	—	—	—
A5. Which session gave you more headache?	—	—	—
Preference of configuration			
C1. Which session did you prefer?	—	—	—
C2. How clear is your 3D sensation?	—	—	—

4.2 Results and Analysis

In order to analyze the statistical significance of the results for the methods used in the comparison, the non-parametric Friedman test [27] was applied. This allows us to measure the statistical significance of the results obtained from the experiments. Tables 3, 4 and 5 show the obtained results. Table 3 analyzes the users' data related to the playability, Table 4 the data concerning the annoyance and Table 5 synthesizes the data regarding the degree of preference. In order to perform a better analysis, they were grouped in rankings, then the number of methods or freedom degrees took a value equal to 3. The configuration valued with lower score by the user had a rank of 1, while the highest score had a rank of 3. In the case where two or more configurations had the same value, an average of the ranks was assigned to them.

The Friedman test considers all user scores to be equal in terms of importance. We used the chi-square distribution to discern the statistical significance of the results. The value of the chi-square distribution in this case was 7.82, with a level of confidence $p = 0.05$, indicating that values lower than this number would imply that there is no statistical significance. Tables 3, 4 and 5 show the statistical significance being positive (+) when the value is higher than 7.82, and negative (-) otherwise.

The first three columns of all the tables show the means obtained from each of the methods for every question that appears in the questionnaire. The standard deviation of the data is shown in parentheses. The results from the statistical Friedman test are shown in the last column. Finally, for a better analysis of the data, the means of the obtained results for every configuration are shown in the last row of the tables.

Analyzing the obtained data, it can be stated that the players prefer the POA configuration in all the analyzed fields. The DCOA is the configuration that occupies the second position in the ranking and finally, the SCOA has been the camera configuration that users have marked as the worst one.

Table 3 shows the data concerning to the playability. The POA is the configuration chosen by the users, with a mean of 3.73. In this case, only the question labeled P6 has conclusive results. It confirms that players enjoy playing more when this configuration is visualized. Results regarding question P5 are the next closest to the value of the chi-square distribution. Then, it can be said that players have felt immerse in this video game. However, the rest of Friedman test values indicates that the results obtained for the other questions are not significant. This is due to the fact that questions are aimed at evaluating the playability of the video game, more than their feelings about the stereo vision. For instance, to evaluate if the video game seems more unpredictable in one configuration than in other (P1) is difficult for the users. They do not find significant differences between the different camera settings when the playability is being compared.

Table 3 Table analyzing users' playability.

Playability	POA	SCOA	DCOA	Friedman
P1	3.20 (± 1.24)	3.40 (± 1.21)	3.40 (± 1.21)	1.8 (-)
P2	3.98 (± 1.03)	3.84 (± 1.09)	3.84 (± 1.00)	0.6 (-)
P3	4.07 (± 0.94)	3.64 (± 1.28)	3.87 (± 1.14)	2.7 (-)
P4	4.31 (± 0.85)	4.11 (± 1.09)	4.20 (± 1.01)	0.6 (-)
P5	3.58 (± 0.99)	3.07 (± 0.91)	3.24 (± 1.11)	5.3 (-)
P6	3.27 (± 1.01)	3.00 (± 0.93)	3.16 (± 1.09)	10.7 (+)
<i>Means</i>	3.73	3.48	3.57	

Regarding the annoyance felt during the visualization (Table 4), the POA configuration also is the one selected in every question. The mean, 4.26, indicates that this configuration is the one that less discomfort produces to the players. Results obtained for questions A3, A4 and A5 have statistical significance and endorse this analysis. In all of them, POA configuration has been the indicated as the one that less headache, eyestrain or fatigue produces. Results obtained for A1 and A2 questions are not significant. Although POA is the option chosen by the players also in these cases, results are not conclusive.

Finally, an analysis of the QoE according to the preferences of configuration (Table 5) reconfirms that users tend to choose the POA method, being the results significantly correlated. The mean for this option (3.85) remains high with respect to the other two, so this fact clearly differentiates it. These results are quite significant because, the means obtained for the POA in both questions, C1 and C2, are really higher than the obtained for the other configurations. Users have indicated that the 3D sensation is better when this configuration is set up (C2), and this is the one selected as favorite for playing in the MVATV (C1).

Table 4 Table analyzing the users' annoyance.

Annoyance	POA	SCOA	DCOA	Friedman
A1	4.22 (± 1.06)	3.98 (± 1.14)	4.13 (± 1.24)	0.5 (-)
A2	4.89 (± 0.32)	4.80 (± 0.59)	4.84 (± 0.56)	0.2 (-)
A3	4.49 (± 0.92)	4.09 (± 1.14)	4.33 (± 1.02)	10.5 (+)
A4	3.60 (± 1.12)	2.91 (± 1.29)	3.40 (± 1.19)	8.7 (+)
A5	4.11 (± 1.17)	3.69 (± 1.18)	4.02 (± 1.27)	9.6 (+)
<i>Means</i>	4.26	3.89	4.14	

Table 5 Table analyzing the users' preference of configuration.

Preference of configuration	POA	SCOA	DCOA	Friedman
C1	3.87 (± 0.89)	2.89 (± 0.96)	3.38 (± 1.03)	13.1 (+)
C2	3.84 (± 0.93)	2.96 (± 1.17)	3.40 (± 1.21)	16.5 (+)
<i>Means</i>	3.85	2.92	3.39	

Regarding the initial hypothesis that appears in Section 2.3, it seemed coherent to think that maintaining the static plane (SCOA) reduces the adverse effects of stereoscopy. However, given the results, the trend indicates that what matters to users is that the most of the objects in the video game appear sharp. So, analyzing the results, it can be concluded that the more objects are focused, the less annoyance is produced. This makes the option POA the most pleasant and the one that offers best quality when a video game is played in a 3D MVATV.

5 Conclusions and Future Work

In this work, different areas of research related to autostereoscopic monitors have been analyzed by focusing the study on the visualization of scenes with movement, such as videogames. Concepts such as the convergence-accommodation conflict and the effect of movement on human visualization have been considered in the analysis. For this purpose, an autostereoscopic rail shooter videogame was designed using a 3D engine and it was visualized on a 3D MVATV that has eight views. Three camera configurations were also proposed in an attempt to prevent users' discomfort: parallel optical axes (POA), static converging optical axes (SCOA) or dynamic converging optical axes (DCOA).

Results reveal that the POA configuration cause less discomfort than the SCOA DCOA. Users that usually play video games are used to controlling all the possible dangers that threaten their avatar. For this reason, they observe carefully all the elements present in the scene. Due to the fact that the comfortable viewing zone in the POA set up is deeper than in the SCOA and DCOA configurations, moving objects are more time sharpen and this makes users mark it as their favorite.

The DCOA configuration is the next best rated. This camera set up has been mainly selected by users that do not usually play. They are concentrated the most of the time in the spacecraft. So, they have marked this option because the focus is all the time on it during the game, so it is always sharp. However, there is a risk of clarity being lost in the visualization of the other objects in the scene. The SCOA configuration has a more reduced comfortable viewing zone that does not move with the spacecraft. The objects can be out of this area more frequently, so they will be blurry more often. This option has been the worst valued by the users.

Concerning the playability, it has been demonstrated that users consider it difficult to choose one configuration over others when evaluating this concept. Only when they are asked about their feelings and impressions, their responses about the chosen configuration is statistically significant.

This work opens up new areas of research on autostereoscopic devices: the user influence of the interface, the period of visualization, the complexity of the game at the design level, the users' experience with different environments and game types, and so on. In addition, it is important to establish a consensus regarding design guides that exploit the capacity of autostereoscopic videogames. Lot of topics related to a field that is having a higher role every day in the video game industry.

Acknowledgements This work was supported by the Spanish Ministry of Science and Technology (Project TIN2016-75866-C3-1-R).

References

1. Tam, W. J.; Speranza, F.; Vázquez, C.; Renaud, R.; Hur, N.: Visual comfort: stereoscopic objects moving in the horizontal and mid-sagittal planes, in: Proc. SPIE 8288, Stereoscopic Displays and Applications XXIII, 82-88(13), (2012)
2. Yano, S.; Emoto, M.; Mitsuhashi, T.: Two factors in visual fatigue caused by stereoscopic HDTV images, *Displays* (25)4, pp. 141-150 (2004)
3. Speranza, F.; Tam, W.; Renaud, R.; Hur, N.: Effect of disparity and motion on visual comfort of stereoscopic images, in: Proceedings of SPIE Stereoscopic Displays and Virtual Reality Systems, 6055, pp. 94-103, (2006)
4. Li, J.; Barkowsky, M.; Callet, P. L.: Visual discomfort of stereoscopic 3d videos: Influence of 3d motion, *Displays* 35(1) pp. 49-57, (2014)
5. Kuze, J.; Ukai, K.: Subjective evaluation of visual fatigue caused by motion images, *Displays* 29 (2), pp. 159-166, (2008)
6. Lambooi, M.; Fortuin, M.; Ijsselstein, W.; Evans, B. Heynderickx, I.: Measuring visual fatigue and visual discomfort associated with 3-d displays, *Journal of the Society for Information Display* 18(11) pp. 931-943, (2010)
7. Woods, A.: Understanding crosstalk in stereoscopic displays, In *International Conference on 3D systems and Applications (3DSA)*, Tokyo, Japan, (2010)
8. Brunnström, K.; Wang, K.; Tavakoli, S.; Andrén, B.: Symptoms analysis of 3D TV viewing based on Simulator Sickness Questionnaires, Quality and User Experience, (2)1, pp. 1-15, (2017)
9. Johnson, D.M.: Introduction to and review of simulator sickness research, Research Report 1832, U.S Army Research Institute, (2005)
10. Hoffman, D.M.; Girshick, A.R.; Akeley, K.; Banks, M.S.: Vergence-accommodation conflicts hinder visual performance and cause visual fatigue, *Journal of Vision* 8(3) pp. 1-30 (2008)
11. Charman, W.N.: The eye in focus: accommodation and presbyopia, *Clinical and Experimental Optometry* 91(3) pp. 207-225, (2008)
12. Howarth, P.A.: Potential hazards of viewing 3-d stereoscopic television, cinema and computer games: a review, *Ophthalmic and Physiological Optics* 31 pp. 115-122, (2011)
13. Holliman, N.: 3D Display Systems. *Science* 38(8) pp. 31-36, (2005)
14. Ukai, K.; Howarth, P.A.: Visual fatigue caused by viewing stereoscopic motion images: Background, theories, and observations, *Displays* 29(2) pp. 106-116, (2008)
15. Li, J.; Barkowsky, M.; Le Callet, P.: The Influence of Relative Disparity and Planar Motion Velocity on Visual Discomfort of Stereoscopic Videos, In: *International Workshop on Quality of Multimedia Experience QoMEX*, Mechelen, Belgium pp.1-6, (2011)
16. Chen, C.; Wang, J.; Li, K.; Liu, Y.; Chen, X.: Visual fatigue caused by watching 3d tv: an fmri study. *BioMedical Engineering OnLine* 14(1) (Suppl 1): S12, (2015)

17. Hassenzuhl, M.; Tractinsky, N.: User experience - a research agenda. *Behaviour & Information Technology* 25(2) pp. 91-97, (2006)
18. Jumisko-Pyykkö, S.; Weitzel, M.; Strohmeier, D.: Designing for user experience: What to expect from mobile 3D tv and video? In: *Proceedings of the 1st International Conference on Designing Interactive User Experiences for TV and Video*, New York, USA, ACM pp. 183-192, (2008)
19. McIntire, J.P.; Havig, P.R.; Geiselman, E.E.: Stereoscopic 3d displays and human performance: A comprehensive review, *Displays* 35(1) pp. 18-26, (2014)
20. Takatalo, J.; Kawai, T.; Kaistinen, J.; Nyman, G.; Häkkinen, J.: User experience in 3d stereoscopic games. *Media Psychology* 14(4) pp. 387-414, (2011)
21. Mahoney, N.; Oikonomou, A.; Wilson, D.: Stereoscopic 3d in videogames: A review of current design practices and challenges, In: *16th International Conference on Computer Games (CGAMES)*, pp. 148-155, (2011)
22. Schild, J.; Masuch, M.: Fundamentals of stereoscopic 3d game design. In: *Proceedings of the 10th International Conference on Entertainment Computing, ICEC11*, Berlin, Heidelberg, Springer-Verlag pp. 155-160, (2011)
23. Jin, Z.X.; Zhang, Y.J.; Wang, X.; Plocher, T.: Evaluating the Usability of an Auto-stereoscopic Display. *Springer Berlin Heidelberg*, Berlin, Heidelberg 18 pp. 605-614, (2007)
24. González, C.; Martínez Sotoca, J.; Pla, F.; Chover, M.: Synthetic content generation for autostereoscopic displays. *Multimedia Tools and Applications* 72(1) pp. 385-415, (2014)
25. Onural, L.; Ozaktas, H.M *Three-dimensional Television: From Science-fiction to Reality*, Springer Berlin Heidelberg, Berlin, Heidelberg pp. 1-9, (2008)
26. Unity 3d game engine, <https://unity3d.com/es> Accessed: 2018-02-02.
27. García, S.; Fernández, A.; Luengo, J.; Herrera, F.: Advanced nonparametric tests for multiple comparisons in the design of experiments in computational intelligence and data mining: Experimental analysis of power. *Information Sciences, Special Issue on Intelligent Distributed Information Systems*, 180 (10) pp. 2044-2064, (2010)
28. Sexton, I.; Surman, P.: Stereoscopic and autostereoscopic display systems. *IEEE Signal Processing Magazine* 16(3) pp. 85-99, (1999)
29. Lambooi, M.T.M., Ijsselsteijn, W.A., Heynderickx, I.: Visual Discomfort in Stereoscopic Displays: A Review. In: *Proceedings of SPIE*, vol. 6490, pp. 649001-64900131, (2007)
30. Brockmyer, J. H.; Fox, C. M.; Curtiss, K. A.; McBroom, E.; Burkhardt, K. M.; Pidruzny, J. N.: The development of the Game Engagement Questionnaire: A measure of engagement in videogame-playing. *Journal of Experimental Social Psychology* 45 pp. 624-634 (2009)
31. Nacke, L.; Lindley, C.A.: Flow and immersion in first-person shooters: measuring the player's gameplay experience. In *Proceedings of the 2008 Conference on Future Play: Research, Play, Share (Future Play '08)*. ACM, New York, NY, USA, pp. 81-88, (2008).
32. Shibata, T.; Kim, J.; Hoffman, D. M.; Banks, M. S.: The zone of comfort: Predicting visual discomfort with stereo displays. *Journal of vision*, 11(8) (2011)
33. Code and data of the developed experiments, <https://github.com/ceviUJI/autostereo>. Accessed: April 2019.