Subjective Evaluation of Super Multiview Video in Consumer 3D Displays

Pablo Carballeira, Jesús Gutiérrez, Francisco Morán, Julián Cabrera, and Narciso García

Abstract—We present preliminary experiments on subjective evaluation of Super Multiview Video (SMV) in stereoscopic and auto-stereoscopic displays. SMV displays require a large number of views (typically 80 or more), but are not yet widely available. Subjective evaluation in legacy displays, though not optimal, will therefore be necessary for the development SMV video technologies. This has lead us to perform standardized subjective evaluation of uncompressed SMV test sequences, simulating SMV displays through view sweep, which is controlled by three parameters: View-Sweep Speed (VSS), Viewing Range, and View Density (VD). In our analysis we have identified ranges of most comfortable values of VSS and VD, providing a comfortable view sweep with smooth transition between views.

I. INTRODUCTION

In the past years, a great development of 3D display technology has taken place, leading to the arrival of stereoscopic and even auto-stereoscopic displays in the consumer market. However, 3D video systems have not achieved the total acceptance of the consumers, due to issues like the lack of high quality 3D video content, and the need of using specific glasses to watch 3D videos in stereoscopic displays, which are the most common 3D monitors nowadays. Moreover, auto-stereoscopic display technology, which is for the moment the most viable alternative to avoid glasses to watch 3D content, does not provide a satisfactory Quality of Experience (QoE) for its price, because of the inherent loss of resolution that severely affects the image quality. Therefore, the intense research work in 3D display technologies has continued over these years [1], and systems such as holographic, integral imaging systems, and SMV displays are being studied and developed towards the possibility of offering a high 3D QoE to the viewers.

In particular, SMV displays seem to be the most promising glasses-free 3D visualization technology, since they offer less technical bottlenecks than, for instance holographic displays [2]. As any multiview visualization system, SMV displays generate a set of discrete views and distribute them over the viewing field in front of the screen, offering a series of viewing zones where the observers can perceive motion parallax as they move their heads, seeing different perspectives of the scene [3]. However, conventional multiview 3D displays have three important drawbacks. Firstly, the effective image resolution is divided by the number of views provided by the display.

Secondly, there is always a discontinuity in view switching with respect to the viewing direction. Finally, they suffer from the accommodation-vergence conflict (the viewer's eyes focus on the screen while they converge in the plane where the objects are projected), which may cause visual discomfort [1]. To overcome these problems, researchers are working on the development of SMV displays that provide sufficient views for continuous motion parallax and for a comfortable viewing experience, since the distance between two adjacent views becomes smaller than the pupil size satisfying the SMV condition [4]. To achieve this requirement, SMV displays require a large number of views, typically 80 or more [5]. In addition, to obtain a satisfactory image resolution, different technologies are being investigated, such as multi-projection systems or flat-panel systems [4].

These types of displays are still under research, so they are not yet widely available and only a few prototypes have been presented [6]. However, there is a need of evaluating the impact on the viewers' QoE of certain effects related to SMV video visualization, such as motion-parallax smoothness or visual comfort. In fact, although only a few studies have been already presented, there are some examples in which the smoothness has been evaluated using a head-up display [7] and the visual comfort effects have been simulated [8]. Furthermore, to obtain a reliable evaluation of the influence of these factors that new 3D displays present, a revision of the traditional methods should be carried out [9].

The current work presents a preliminary study of the impact of various factors on the subjective evaluation of SMV content using legacy displays. We have taken as reference a similar study that was proposed for the previous generation of multiview displays [10]. The functionality of SMV displays has been simulated implementing a view sweeping of SMV sequences, which allowed the evaluation of the effects of the View-Sweep Speed (VSS), the View Range (VR), and the View Density (VD). The objective was to derive ranges of comfortable VSS and VD, and determine the influence of the view-sweep effect and the amplitude of the VR in the 3D perception of the scene. The results and conclusions of the current work may provide relevant inputs for future subjective evaluations of SMV content [11] and for standardization purposes within MPEG's FTV community [12].

The rest of the paper is structured as follows. In Section II a detailed description of the simulation of SMV video
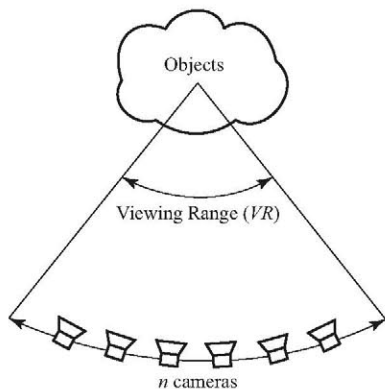
Fig. 1: SMV scenario

Fig. 2: Diagram of the ACR method

visualization with view sweeping is provided. In Section III, the specific configuration of the subjective test is described and the obtained results are presented in Section IV. Finally, Section V provides some general conclusions.

II. SMV DISPLAY SIMULATION AND TEST CONTENT

The objectives of SMV displays is to provide a viewing experience without visual discomfort and a smooth motion parallax as the observers move their heads to see different viewpoints of the scene. To achieve this, the images projected by the display should not be so different between neighboring viewpoints that the observers could perceive discontinuities. This difference between neighboring views could be affected by the disparity of the content, the separation between the cameras when capturing the scene, and the rate of displaying new viewpoints for a certain head movement [10]. Thus, this motivates the research on the influence of the view density and speed of head movements on the QoE with SMV.

Due to the unavailability of SMV displays, since only a few prototypes have been developed, these parameters should be investigated using common 3D displays simulating the head movement of the observers to see different views of the scene by displaying a sweep of those different views to static observers. The following subsections describe the technical details of this simulation to achieve realistic test conditions.

A. View-sweep parameters

Figure 1 depicts a typical SMV scenario, in which the scene is captured by a set of n cameras, typically 80 or more, original or synthesized, spanning a certain total angle *VR*. The user navigation changing the viewpoint within *VR* has been simulated by means of view sweep, i.e. by composing a video sequence whose (concatenated) frames are consecutive views. The parameters that define the view sweep, other than n and *VR*, are the following (where t represents time in seconds):

$$\text{View Density: } VD = \frac{n}{VR} \qquad (1)$$

$$\text{View-Sweep Speed: } VSS = \frac{VR}{t} \qquad (2)$$

Thus, one sequence can be viewed for a given *VR* with a given *VSS* and a given *VD*. The definition of the view-sweep
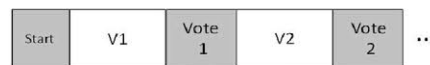
parameters are defined for cameras arranged in an arc in which the angular change between cameras is constant. A linear arrangement of cameras distant enough from the scene object could also be easily approximated by this parametrization.

B. SMV sequences

In this work, we have used the two multiview video sequences with the highest camera density that are currently available in the category of MPEG's SMV test material [5]: Bee and Shark. Their characteristics are shown in Table I.

Given the camera configuration in the test sequences [13], $VR = 41.62°$. Thus, the maximum *VD* available using original cameras only is $VD_{\max} = 4.42$ cam/°. Note that the cameras are not arranged in an arc, thus, the angular change between consecutive cameras is not constant. However, the average angular change is $0.226°$ while the minimum and maximum angle changes are $0.207°$ and $0.237°$.

III. SUBJECTIVE TEST SETUP

A. Methodology

The Absolute Category Rating (ACR) method was used to evaluate the test sequences. Thus, as depicted in Fig. 2, after an initial message indicating the start of the test, the test sequences (*Vi*) were shown followed by a message (Vote *i*) indicating the observers to rate the corresponding video sequence. To collect the opinions of the observers, questionnaires with numbered boxes were used, where they were asked to write a mark for the corresponding evaluation.

The voting messages had a duration of 10 seconds, to allow the subjects to judge the following factors after watching each test clip:

- Smoothness of the view transitions using a 5-grade quality scale [14].
- 3D quality using the 5-grade quality scale [14] and taking into account factors such as depth perception, immersiveness, etc.
- Visual comfort using the 5-grade comfort scale [15] and considering all the factors that could make uncomfortable the visual experience, such as difficulties in focusing and converging, visual discomfort caused by depth, comfort of the view sweeping speed, etc.

The test sessions consisted of a previous visual screening of the subjects, followed by a training process in which some example sequences were shown to them explaining the tests methodology and purpose. This also provided a reference to the observers of the 3D capabilities of the displays, so they could rate the factors to evaluate (especially 3D quality) taking into account the performance of each display and without directly comparing them (which was not the purpose of the tests). Then, the test videos were shown to the observers in each of the displays, with a rest period of approximately five

TABLE I: CHARACTERISTICS OF THE TWO MPEG SMV TEST SEQUENCES USED

	Resolution	Frame rate	Movement	Number of cameras	Camera distance	Angle between adjacent cameras
Bee	1920x1088	-	Static	185	3.711 mm	˜0.226°
Shark	1920x1088	30 fps	Moving	185	7.004 mm	˜0.226°

minutes during the display change. Different randomizations of the videos were used in each session and for each display to reduce contextual effects, with the condition of showing twice each test sequence but never consecutively. In addition, the display order was changed in each session. Two observers participated at a time in each session, which lasted one hour approximately.

B. Test material

For the subjective tests, we have prepared test sequences with three different scene contents:

- Bee (static).
- Shark static: obtained using frame 68 of the sequence.
- Shark motion: using all 100 frames of the sequence.

Regarding the simulation of the SMV display through view sweep, we have tested using the full VR (41.62°) and half VR (20.81°). Each sequence was displayed at five different $VSSs$: 3.39, 6.78, 13.56, 27.12 and 54.24 °/s. For each value of VSS, we selected four different values of VD (see Table II). Those VSS and VD values are coherent with previous similar works [10]. Additionally, also a fixed-viewpoint sequence (VSS_0) was shown for each case. Given that all test sequences have the same length, and are shown at different values of VSS, the number of times that the camera covers VR is different for each value of VSS. All tests sequences last 10 seconds, and were displayed at a spatial resolution of 1920x1088 pixels (in side-by-side format) and a frame rate of 60 fps. In the case of shark motion, the frame rate was converted from 30 fps to 60 fps by frame repetition with a factor of x2. The stereo baseline for all test sequences was set to 5.57 cm.

C. Environment

The test area was set according to international recommendations [14], with walls covered by mid-gray curtains, and the ambient lightning conditions were controlled to avoid disturbing reflections. The viewing distance was set at 2.1 m from the display position, which corresponds to a viewing distance of 3H and it is within the viewing distance ranges of the displays used.

D. Equipment

Two displays (one stereoscopic and one auto-stereoscopic) were used to carry out the subjective tests, both 55", in order to avoid the influence of screen size in the visual experience: a curved UHD LED consumer stereoscopic display from Samsung (model UE55HU8500L; release year: 2014) with active shutter glasses (model SSG-5100GB), and an auto-stereoscopic display from Toshiba (model 55ZL2G; release year: 2012). The latter is based on lenticular technology to provide nine views (taking a stereo pair as input and using a proprietary algorithm to create seven more perspectives)

with an effective resolution of 720p, thanks to a total panel resolution of 3840x2160 pixels and a pixel layout to group every nine pixels. Given the importance of the viewing position of the observers to watch 3D content in autosteroecopic displays, the TV set was calibrated, and the correct position of the observers was assured using its own tracking system. It is worth noting that, in practice, the viewing cone of the auto-stereoscopic display is very narrow, providing the perception of a single viewpoint.

The videos were played using a PC with a graphic card Nvidia GeForce GTX 760 and the Media Player Classic Home Cinema (MPC-HC) [16], which allowed a smooth playback of the sequences.

E. Observers

A total of 22 observers (6 females, 16 males) participated in the tests, all of them having normal or corrected vision. The ages of the participants were ranged between 22 and 49, with an average age of 30. Furthermore, a screening of the results scores provided by the observers was carried out analyzing the Pearson correlation between the data provided by each subject and the average of all the subjects [17], which lead to discard two observers.

The participants, all of them from our university, were asked to fill a questionnaire where they indicated their experience in visualizing 3D video. All of them reported to have watched 3D content before the tests. In particular, 9% of the observers had watched 3D video once, 77% occasionally, and 14% frequently.

IV. EXPERIMENTAL RESULTS

In this section, the main results obtained from the experiment are reported. The results shown in the figures were obtained computing the Mean Opinion Scores (MOS) from the evaluations provided by the observers in the questionnaires. In addition, the 95% Confidence Intervals (CI) are represented according to the computation recommended by ITU-R BT.500 [14].
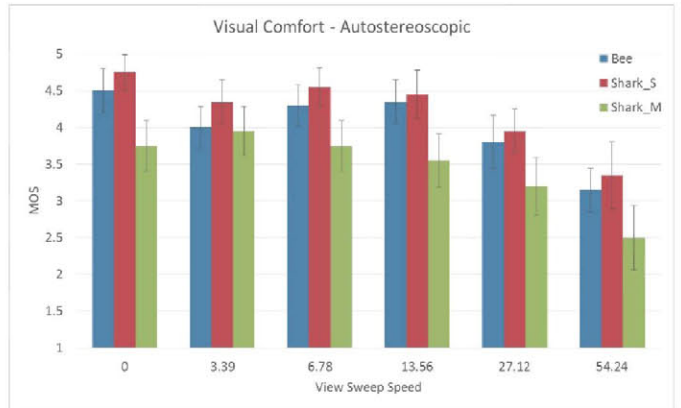
A. Comfortable view-sweep speed

The first objective of the study was to establish a range of VSS values that are comfortable to view SMV content in stereoscopic and auto-stereoscopic displays. Therefore, we have analyzed the visual comfort results, considering that a VSS is comfortable if the MOS score of the visual comfort is strictly above 3 (including the entire range of the CI). Fig. 3a and Fig. 3b show the MOS results for the visual comfort for all values of VSS for the stereoscopic and auto-stereoscopic displays respectively. Concretely, for each value of VSS, the results for VD_1 are shown (i.e. the highest view density that has been tested). As a reference, we also include the visual

TABLE II: COMBINATIONS OF *VSS* AND *VD* TESTED

	VSS_0 0 °/s	VSS_1 3.39 °/s	VSS_2 6.78 °/s	VSS_3 13.56 °/s	VSS_4 27.12 °/s	VSS_5 54.24 °/s
VD_1 (cam/°)	-	4.42	4.42	4.42	2.21	1.11
VD_2 (cam/°)	-	2.21	2.21	1.11	0.74	-
VD_3 (cam/°)	-	1.47	1.11	0.74	0.44	-
VD_4 (cam/°)	-	1.11	0.74	0.44	0.28	-
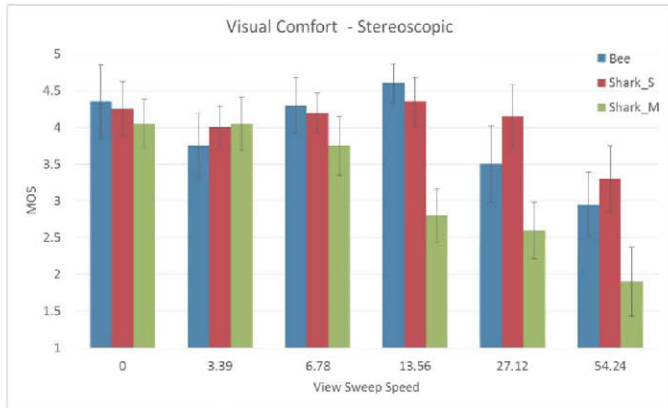
(a) Stereoscopic display

(b) Visual Comfort. Auto-stereoscopic display

Fig. 3: Visual comfort results for different values of *VSS*. Results shown for each test sequence: *Bee*, *Shark static* and *Shark motion*

TABLE III: COMFORTABLE VALUES OF VIEW-SWEEP SPEED

		Highest comfortable *VSS*	Most comfortable *VSS* range
Stereo	Static	13.56 °/s	[6.78 − 13.56] °/s
	Moving	6.78 °/s	[6.78 − 13.56] °/s
Auto	Static	27.12 °/s	[3.39 − 6.78] °/s
	Moving	13.56 °/s	[3.39 − 6.78] °/s

comfort results for the fixed-camera sequences ($VSS_0 = 0$). In addition, Table III shows the highest value of comfortable *VSS* and the most comfortable *VSS* range for each type of content and display.

As shown in the figures, the visual comfort generally decreases with increasing *VSS*, except for the lowest *VSS* in static-content sequences, for which the highest *VD* values are not enough to provide a smooth camera navigation, affecting the observers' comfort. Moreover, comparing both types of content, the static scenes allow higher values of *VSS* as the combination of scene motion and camera movement creates visual discomfort. Finally, regarding the display, the auto-stereoscopic provides better comfort, even for high *VSS*, probably thanks to not needing glasses to watch 3D content.

B. Comfortable view density

The second objective was to find the range of values of *VD* that provide a perception of smooth transition through consecutive views. To find those values, the smoothness scores indicated by the observers were analyzed, considering again that the view sweep is smooth enough if the MOS score of the smoothness is strictly above 3 (including the entire range

of the CI). Fig. 4a and Fig. 4b show the MOS results for the smoothness for all values of *VSS*, in the stereoscopic display, for the static and moving content respectively. Fig. 4c and Fig. 4d show the same type of results, in the auto-stereoscopic display, for the static and moving content respectively. Note that for VSS_5 only one value of *VD* was tested.

As we hypothesized, these ranges of "smooth" *VD* values depend on *VSS*. Thus, in Table IV we provide the minimum value of *VD* that is acceptable for a given value of *VSS*. We are especially interested in the *VSS* values that were identified in Section IV-A as most comfortable for each type of content and display. Those results are marked in blue.

The main conclusions that can be drawn from the analysis of the smoothness results are the following:

- As stated in Section IV-A, high values of *VSS* in the moving sequence are uncomfortable, and this discomfort may have influenced the observers when ranking the smoothness of the view sweep, leading to irrelevant results (marked in red in Table IV).
- Regarding the differences between both types of content for the stereoscopic display, the moving sequence can be shown with lower *VD*, and be still perceived as smooth.
- For the auto-stereoscopic display there is no remarkable difference in the values of minimum *VD* that are smooth for both types of content (for comfortable *VSS* values).

C. Effect of view sweep in 3D perception

Another objective of this work was to analyze the influence of the view-sweep effect in the 3D perception of the observers, i.e. perception of depth and immersivity. Thus, we compare the
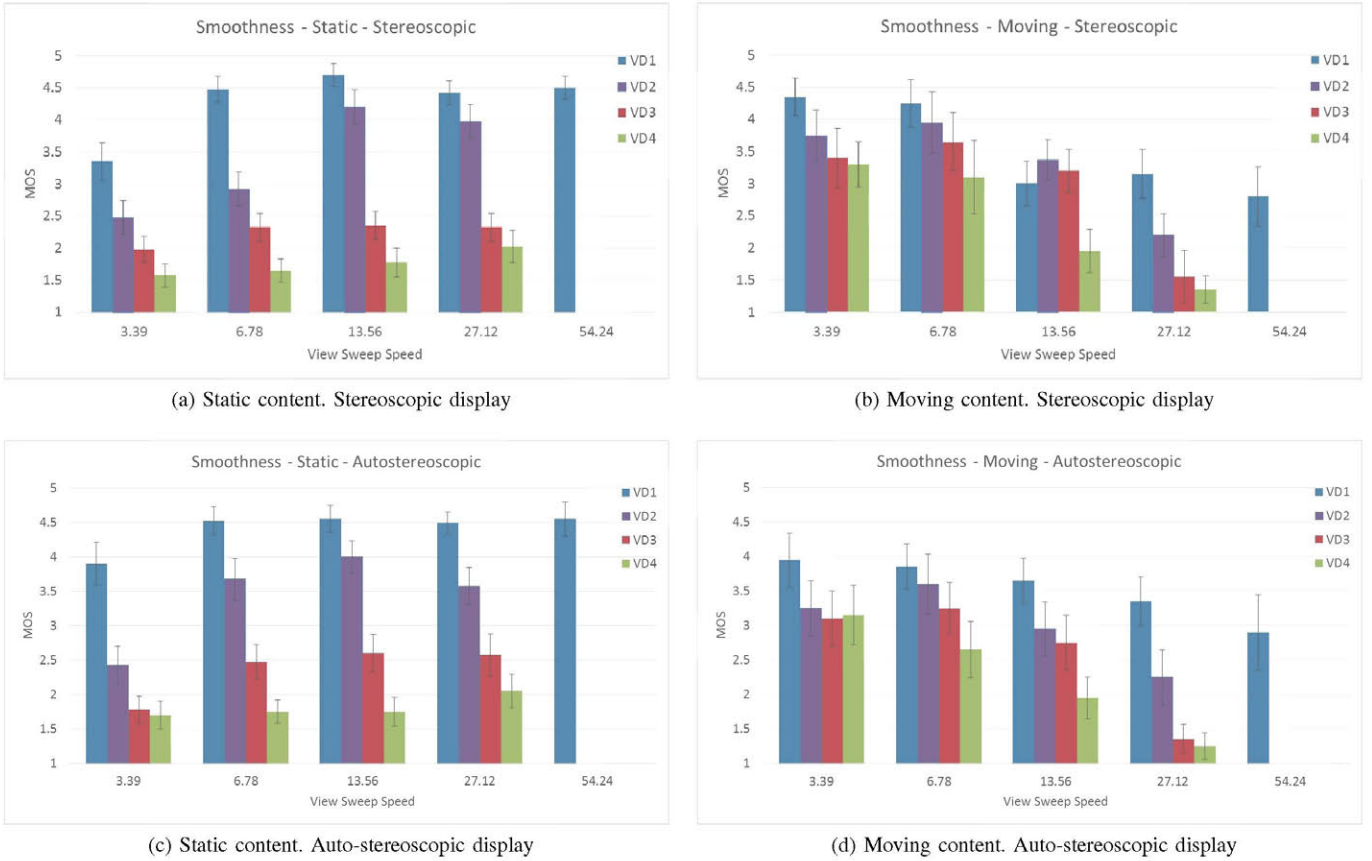
(a) Static content. Stereoscopic display

(b) Moving content. Stereoscopic display

(c) Static content. Auto-stereoscopic display

(d) Moving content. Auto-stereoscopic display

Fig. 4: Smoothness results for all the combinations of *VSS* and *VD*. Each group of columns corresponds to one value of *VSS*. Within each group, each column corresponds to a different value of *VD*. The *VD* values tested for each *VSS* are shown in Table II. Results averaged for ech type of scene content: static and moving.

TABLE IV: MINIMUM VALUE OF *VD* PER VALUE OF *VSS*. RESULTS MARKED IN BLUE AND RED CORRESPOND TO: THE MOST COMFORTABLE *VSS* AND UNCOMFORTABLE *VSS* VALUES, RESPECTIVELY, FOR A GIVEN CONTENT TYPE AND DISPLAY. ALL *VD* RESULTS ARE GIVEN IN CAM/°

	Stereo		Auto	
	Static	Moving	Static	Moving
VSS_0	-	-	-	-
VSS_1	X	$VD_2 = 2.21$	$VD_1 = 4.42$	$VD_1 = 4.42$
VSS_2	$VD_1 = 4.42$	$VD_3 = 1.11$	$VD_2 = 2.21$	$VD_2 = 2.21$
VSS_3	$VD_2 = 1.11$	$VD_3 = 0.74$	$VD_2 = 1.11$	$VD_1 = 4.42$
VSS_4	$VD_2 = 0.74$	X	$VD_2 = 0.74$	$VD_1 = 4.42$
VSS_5	-	-	-	-

3D quality of the view-sweep sequences for different values of *VSS* against the fixed-camera sequences.

Fig. 5a and Fig. 5b show the MOS of the 3D quality results for each different content and *VSS* with the maximum *VD* (VD_1), to avoid the influence of the lack smoothness in the analysis. The main conclusions that we draw form these results are the following:

- The 3D perception improves with view sweep for static sequences, but not for the moving sequences. This im-

provement is higher for auto-stereoscopic displays, in which the depth perception is lower initially.
- This 3D perception drops for sequences with motion at high *VSSs* that are not comfortable (see Section IV-A).

D. Influence of the viewing range

Finally, we have also analyzed the influence of the amplitude of *VR* in the visual comfort and the 3D quality perceived by the observers. The results showed that:

- There is no statistically significant difference in visual comfort between the MOS scores for full and half *VR* with the sequences considered in the test.
- Although the scores for 3D quality with full *VR* are slightly higher, the difference between the MOS scores for full and half *VR* are not statistically significant.

E. Other results

The participants were asked to fill a questionnaire after the test in which they were asked whether they felt visual fatigue or headache. Concretely, the fatigue was measured considering the five-grade scale defined in [18] (ranged from: 5-"My eyes are not tired" to 1-"My eyes are tired"), and a similar one was used for headache. It is worth noting that almost 55% of

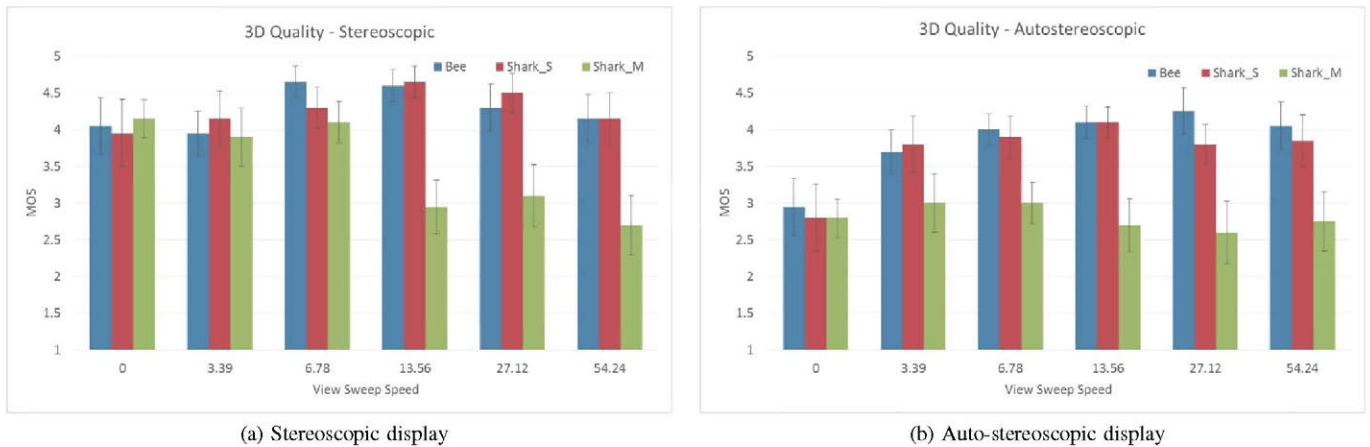(a) Stereoscopic display

(b) Auto-stereoscopic display

Fig. 5: 3D quality for the maximum *VD*. Results shown for each test sequence: *Bee*, *Shark static* and *Shark motion*

the observers expressed that they felt visual fatigue, and 23% reported a slight headache.

V. CONCLUSIONS

SMV video technology is currently the most promising alternative for a future glasses-free high-quality 3D experience. Despite the fact that the development of these display kinds is under research, so only a few prototypes have been already presented, it is necessary to study the impact on the viewers' QoE of the new effects that arise with SMV displays, such as view density and visual comfort. Therefore, preliminary approaches analyzing these effects with available 3D displays simulating the operation of SMV systems are crucial for an appropriate development of this technology.

To cover this necessity, this paper presents a novel subjective study of SMV video, simulating the functionality of SMV displays, through view sweep using a stereoscopic and an auto-stereoscopic display. This way, it is possible to analyze the effects of *VD*, *VSS*, and *VR*, evaluating factors such as smoothness, visual comfort and depth perception.

These initial experiments are limited due to the lack of content variety, especially in the case of moving sequences, and the number of values of *VSS* and *VD* that can be subjectively tested. Despite these limitations, the results and conclusions of this innovative study can serve as an initial guide for setting up subjective tests, and refined with further experiments. For instance, the most comfortable *VSS* for static content is around [6.78-13.56] °/s, while for moving content it is more restricted, around [3.39-6.78] °/s. Moreover, the current study has allowed to obtain the minimum *VD* values for smooth view sweep, which are shown in Table IV of Section IV-B. Finally, the influence of the viewing angle seems to be less important than the other parameters studied.

Further studies are being carried out and these conclusions will be extended with new content and a deeper analysis of the of these and other factors related to SMV video, such as coding impairments.

REFERENCES

[1] J. Geng, "Three-dimensional display technologies," *Advances in Optics and Photonics*, vol. 5, no. 4, pp. 456–535, Dec. 2013.

[2] P. Surman and X. W. Sun, "Towards the reality of 3D imaging and display," in *Proc. of IEEE 3DTV Conference*, Budapest, Hungary, Jul. 2014, pp. 1–4.

[3] S. Reichelt, R. Haussler, G. Fütterer, and N. Leister, "Depth cues in human visual perception and their realization in 3D displays," in *SPIE Three-Dimensional Imaging, Visualization, and Display*, vol. 76900B, Orlando, USA, Apr. 2010.

[4] Y. Takaki, "Development of Super Multi-View Displays," *ITE Trans. on Media Technology and Applications*, vol. 2, no. 1, pp. 8–14, Jan. 2014.

[5] ISO/IEC SC29WG11, "Experimental Framework for FTV," Output doc. N15048, 110th MPEG meeting, Strasbourg, France, Oct. 2014.

[6] M. Tanimoto, "FTV standardization in MPEG," in *Proc. of IEEE 3DTV Conference*, Budapest, Hungary, Jul. 2014, pp. 1–4.

[7] Y. Takaki, Y. Urano, and H. Nishio, "Motion-parallax smoothness of short-, medium-, and long-distance 3D image presentation using multi-view displays," *Optics express*, vol. 20, no. 24, pp. 366–375, Nov. 2012.

[8] D.-w. Kim and S.-k. Kim, "Visual Comfort Range in the Super-Multi-View Display," in *International display research conference*, Sao Paulo, Brazil, Nov. 2010, pp. 1–4.

[9] M.-C. Park and S. Mun, "Overview of Measurement Methods for Factors Affecting the Human Visual System in 3D Displays," *Journal of Display Technology*, vol. 9, p. 1, Jan. 2015.

[10] F. Speranza and W. Tam, "Perceived smoothness of viewpoint transition in multi-viewpoint stereoscopic displays," in *Stereosc. Displays and Virtual Reality Systems XII*, vol. 5664, San Jose, USA, Mar. 2005, pp. 72–82.

[11] ISO/IEC JTC1/SC29/WG11, "Draft Call for Evidence on FTV," Output doc. N15095, 111th MPEG meeting, Geneva, CH, Feb. 2015.

[12] P. Carballeira, J. Gutiérrez, and F. Morán, "Subjective evaluation of Super Multiview Video in legacy displays," Contrib. M35823, 111th MPEG meeting, Geneva, CH, Feb. 2015.

[13] T. Senoh, A. Ishikawa, M. Okui, K. Yamamoto, and N. Inoue, "FTV AHG: Super-Multiview Sequences of NICT," Contrib. M32201, 107th MPEG meeting, San Jose, USA, Jan. 2014.

[14] ITU-R, "Methodology for the subjective assessment of the quality of television pictures," Rec. ITU-R BT. 500-13, vol. 13, Jan. 2012.

[15] ——, "Subjective methods for the assessment of stereoscopic 3DTV systems," Recommendation ITU-R BT.2021, Aug. 2012.

[16] Media Player Classic Home Cinema (MPC-IIC), Available at: http://mpc-hc.org/.

[17] ISO/IEC JTC1/SC29/WG11, "Report of subjective test results from the call for proposals on 3D video coding technology," MPEG2011/N12347, Geneva, CH, Geneva, Nov. 2011.

[18] S. Yano, S. Ide, T. Mitsuhashi, and H. Thwaites, "A study of visual fatigue and visual comfort for 3D HDTV/IIDTV images," *Displays*, vol. 23, no. 4, pp. 191–201, Sep. 2002.