PUPILLARY REACTIVITY TO NON-PHOTOREALISTIC RENDERING: A CASE STUDY OF IMMERSION IN 3D CINEMA

Victor FAJNZYLBER (1), Samuel MADARIAGA (2), Rafael DEL VILLAR (1), Pedro MALDONADO (2), Diego Vargas (5), Aria SERRA (6), Milan MAGDICS (4), Mateu SBERT (3,6).

(1) University of Chile, Communication and Imaging Institute.
(2) University of Chile, Departamento de Neurociencia & Biomedical Neuroscience Institute.
(3) Tianjin University, School of Computer Science and Technology.
(4) Budapest University of Technology and Economics, Department of Control Engineering and Information Technology.
(5) National Autonomous University of Mexico, National School of Higher Studies Campus Morelia.
(6) Girona University, Institute of Informatics and Applications & Graphics and Imaging Laboratory.

ABSTRACT

The aim of our research is to build an empirical indicator of cinematic immersion based on ocular data. We study immersion as a gradual and adaptive process of corporal involvement in the cinematic experience. To study its temporal evolution, we analyze correlations between the film and the ocular response, in its attentional aspects (what do we see?) and pupillary aspects (how do we react?). To understand in what ways post-production can favor immersion, we are interested in studying the pupil reactivity to visual attributes of the film: the visual abstraction of a non-photorealistic rendering (NPR) post-production and the optical hyperrealism of stereoscopy. We expect that our results could be useful for conceiving new NPR tools for 3D films and innovative post-production strategies for virtual reality.

Index Terms— Non-photorealistic rendering, pupillary reactivity, film perception, 3D cinema.

Funding: This publication was supported by the National Commission for Scientific and Technological Research by Fondecyt Project No. 1120064 (Chile); Iniciativa Científica Milenio ICM P09-015-F (Chile), the PROA Program of the University of Chile to promote academic productivity (Chile), the Hungarian National Scientific Research Fund OTKA K-124124 project and by Grant TIN2016-75866-CE-3-R from the Spanish Government.

1. INTRODUCTION

In this article, we will explain our research process centered on pupillary reactivity as an indicator of cinematic immersion. In the framework of a case study based on the analysis of ocular data of 27 subjects, we try to answer this question: How can we describe empirically the experience of cinematic immersion? A preliminary definition could be the following: we attend the exhibition of a film and allow ourselves to be carried away by its images and sounds, capturing our attention until it leads us, sometimes, to experience unexpected emotional states. We are interested in understanding cinematic immersion by correlating attentional factors and pupillary responses as manifestations of the corporal involvement during the film.

2. RELATED RESEARCH

In a study about immersion on film viewing [1] it is indicated that the greater the immersion, the greater the emotions that the spectators could feel. They also make the difference between "fictional" immersion and "artifact" immersion, distinguishing the narrative dimension from the visual and sound forms. In other words, a fiction seems to trigger immersion processes that may be convergent or dissonant with the immersive effects triggered by a photographic or lighting style. Sometimes a film is able to capture our attention and involve our body in the experience, for example, when it triggers unexpected emotions. We are interested in this particular kind of immersion: when the sensory experience is accompanied by an emotional response. We know that emotions require time to develop, and for this reason, we are interested in knowing how sensory and cognitive dynamics may be correlated during the temporality of a film, creating an appropriate context for the emergence of emotions. To achieve this, we chose a descriptive strategy that places the temporal evolution of the gaze at the center of the analysis.

To understand the temporal evolution of immersion, we need to describe the film temporality and its dynamic interaction with the behavioral dynamics of the ocular

system during the viewing of a movie. We analyze the pupillary reactivity to evaluate if a visual post-production or a sound design style can trigger corporal involvement of the spectator. We did not use verbal reports or questionnaires filled by spectators. We chose to use ocular behavior as the unique data in our experimental research because it is a measurable object of a corporal involvement linked to vision.

This research was born from the interest of exploring the combination of two cinematic procedures that, at first, seemed almost contradictory: the optical hyperrealism of stereoscopy and the visual abstraction of non-photorealistic rendering (NPR). In a previous study [2] we analyzed the impact of NPR on the perception of 2D cinema, combining viewing with qualitative and quantitative research methods. Viewers reported different reactions to the use of NPR, pointing out a possible correlation between visual abstraction and sound perception: with greater visual abstraction, greater importance was given to the perception of voices as the main mechanism for identifying the characters. However, the most striking result was the idea that the more abstract version of NPR post-production, despite the subtraction of a large amount of visual information from the scene, did not seem to affect the narrative understanding of the film. Both results suggest that the viewer transfers attention resources from vision to hearing, thus modifying the experience of seeing the film, without losing sight of the story and its development. Later we began our second study on NPR and film immersion. We wonder if the same adaptive reaction to visual abstraction observed in a 2D movie could be reproduced with a 3D movie. It seems relevant to us to investigate the cinematic immersion in 2D and 3D, especially in the current context of the growth in the world production of virtual reality films and experiences in augmented reality, grouping both modalities under the denomination of "XR" (extended realities). In the field of immersive contents, the distinction between a 2D and 3D experience is relevant due to its effect on the perception of immersion. Considering this context, we began to study the attentional effects of NPR in the visual perception of 3D cinema. We found these results.

In a cinematic scene shot in stereoscopic 3D, the visual abstraction of the NPR seems to trigger an increase of visual attention on the characters. We wanted to know if this "narrative immersion", based on a preponderant focus on the characters, could be increased through an NPR post-production with a high level of visual abstraction. To test this idea we use ocular behavior data collected with eye-tracking technology. The next step was to describe specifically the attentional and emotional effects of NPR in the perception of 3D cinema. We observe the attentional effects of NPR considering the dynamic behavior of the ocular fixation [3]. This led us to review similar studies in the same field. In a previous study, we found interesting evidence, not necessarily generalizable but highly suggestive: the NPR may capture the viewer's attention with moderate abstraction [4] and with the presence of faces [3]. About the emotional response that could be triggered by the use of NPR processing, it has been found that, with certain formal modifications to the image, the emotional response of the spectator can be guided [4].

It could be thought that the levels of abstraction of a scene would help to achieve this effect, however, if the abstraction effect used is exaggerated or unjustified, it can cause the viewer to be totally detached from the image itself, thus causing the opposite effect to immersion. We find ourselves with a paradox: although visual abstraction does not seem to affect narrative comprehension, it could affect immersion. But in what way? It has been said that the abstract processing of NPR can be useful not only to affirm certain emotions, but also to produce the sensation of being “elsewhere" [5], thus favoring an immersive effect on the viewer. We could feel immersed without being emotionally involved. Movies are not always capable of emotion, but when they do, we assume that we all react differently.

To go further in this duality between immersion and emotion, we are interested in the pupillary reactivity of the spectator, considering that this physiological information has been used as an indicator of emotional response [6].

The pupil (as well as the eye movements and blinking) work as indicators of our internal mental, cognitive and emotional processes, is very useful to recognize certain reactions produced by stimuli or by situations of variable cognitive complexity. In several studies on the pupillary response, different perspectives for studying pupillary reactivity have been described. For example, fluctuations in the diameter of the pupil are a behavioral indicator. A considerable pupil dilation is associated when imagining pleasant or unpleasant situations [6]. Changes in the pupil are greater if the images are more emotionally charged, compared with neutral images, regardless of whether they are pleasant or not [7]. Pupillary reactivity has also been used as an indicator of attentional levels depending on the complexity of the task, as well as the subject's concentration capabilities [8]. In situations of greater analytical and cognitive demands, it seems to be associated with greater pupillary dilation [8]. An example: to follow the thread of a story and simultaneously perceive the optical illusion of a stereoscopic object emerging in 3D from the screen, could be considered as an experience of high cognitive complexity for the intellective action of the viewer. This idea seems relevant to us: the cognitive complexity of the intellective task could be related to pupillary fluctuations. Based on this evidence, we will use the attentional response and the student's pupillary reactivity as the basis for our immersion indicator.

To evaluate its potential usefulness, we studied the possible correlations between the following factors:

- Regardless of the visual style or sound used, in the four versions of the same stereoscopic film we observe variations

in luminosity generated by the camera or people movements. In response to these light changes, we expect an automatic pupillary response.

- We analyze the pupillary variations in each one of the four videos used.

- If the pupillary reactivity behaves in an expected manner due to light fluctuations, we will assume that there is a basis of immersion of the viewer in the film.

- If the behavior of the pupil reactivity does not present global statistical correlation with the temporal development of the luminosity, we will assume that this may reveal a different modality of immersion, eventually correlated with an emotional response to the depicted scene. In order to produce an analytical approach that allows elucidating if there are different immersion modalities, we will analyze possible correlations between pupillary behavior, attentional patterns and the temporal sequence of events that constitute the cinematic narrative.

3. METHODS

We conducted an experiment with a total of 27 anonymous students in the city of Santiago (Chile): 14 women and 13 men, between the ages of 18 and 23, who voluntarily participated in this experiment, using informed consent. Subjects with visual pathologies (myopia, astigmatism, etc.) were excluded to produce comparable ocular responses to audiovisual stimuli. To collect the data, a screening room was prepared with the "The Eye Tribe" eye-tracker mounted below the LG, LED 55 Full HD Smart TV 3D/55LB6500. The eye tracker was used at 30Hz, while the television was configured at a resolution of 1920 * 780 pixels.

3.1. Video processing.

Video selection for the experiment: photorealistic 3D vs. non-photorealistic 3D. We edited a stereoscopic short film of 3 minutes and 30 seconds [9]. This film was processed with a "pastel" filter from an NPR software prototype, extracting color and texture information in order to generate a monochrome non-photorealistic version of the video (black lines on white backgrounds). We selected an NPR filter that can be perceived as a high degree of abstraction. We are aware of the influence of artifacts on NPR perception, as stated by Mould et al. [10] but we chose a not visually appealing NPR style, with visual artifacts, in order to explore a highly abstract aspect very different from the photorealistic sharpness typically used in 3D cinema. We wondered whether this perceptual efficiency associated with NPR could be obtained with a high degree of abstraction, but without affecting the immersive effect of the stereoscopic depth. We produced two videos in NPR mode and two videos in PR mode (photorealistic), example shots are shown in Figure 1. In total, we had four videos of the same duration.

3.1.1. Non-photorealism and visual abstraction.

From a technical point of view, NPR methods can be classified into object-space methods and image-space (also called screen-space) methods [11]. Object-space methods work on the 3D model of the scene and use geometric information in their computations, such as surface normals, curvature, or distance from the camera.

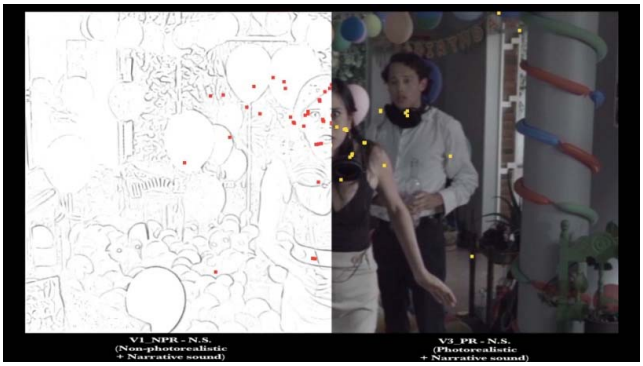

Fig. 1. Styles used in the experiment: non-photorealistic vs. photorealistic.

Many NPR effects operate with randomly adding or removing lines or texture details such as hatching [12] or watercolor rendering [13]. Computing the effects in the 3D object space we can make these random decisions fixed with respect to the object surfaces, resulting in a more coherent look [12] and also an improved stereo coherence [13]. The drawback is that we need to capture or create a complete 3D model of the scene or at least a depth map containing the distances from the camera. Screen-space methods perform their computations on the captured or rendered images and thus generate their output purely based on color information. On the one hand, this makes them easily applicable to any kinds of films. On the other hand, we are limited to effects that have no higher-level knowledge about the scene geometry. In this work, our aim was to use effects that are as general as possible and thus we apply screen-space methods. We note, however, that there are dozens of methods to generate a depth map from stereoscopic images [14], which would allow a limited use of object-space methods. On the other hand, the depth maps generated by these methods often fail to meet the quality requirements of cinematography. There are many NPR algorithms that are designed to create the abstraction on the image [15, 16] and thus allowing to enhance or guide perception [17, 18]. This makes them also suitable for cinematography. However, very few works discuss the issues related to NPR methods in the context of cinematic stereoscopy. Even in the great and very comprehensive survey of stylization techniques by Kyprianidis et al. [19] there is no mention of stereoscopy. Most stereoscopic NPR approaches include line drawings

[20], watercolor rendering [13] or artistic stylization [21], however, we are interested in abstraction methods. As an edge enhancement method, we used the flow-based, extended difference-of-Gaussians (DoG) [22] filters proposed by Winnemöller, as it applies a flow-oriented smoothing step on the edges. Additionally, this method was shown to produce aesthetically pleasing results and can simulate various effects, such as black and white or colored pastel. Abstraction is achieved by a method based on [16]. Texture details are removed by the iterative application of the bilateral filter, whereas color complexity is reduced using a smooth quantization step [22] that avoids sharp transitions and also greatly increases temporal coherence makes the method applicable to videos or other real-time applications such as computer games [23]. Based on these references, for the purpose of our study, we decided to use a real-time processing technique based on a screen-space method, to evaluate the impact of this NPR filter, which seeks to produce a high level of abstraction, in order to induce cognitive uncertainty during the cinematic experience. Two examples of the NPR filter used in the video stimuli can be viewed in [9] and [24].

3.1.2. Sound post-production.

Audio mixing: narrative vs. immersive. Once the visual processing took place, we conducted a process of audio post-production, which consisted of producing two different versions from the same audio material of the short film. The first post-production, deemed "narrative", consisted of emphasizing the voices as the main element of the sound mix. The other post-production, deemed "immersive", put all of the elements of ambient sound at the same level, including the voices of the characters as part of the total sound mix. The "narrative" sound was labeled S.N. and the "immersive" sound as S.I. In total, we generated two videos with narrative sound (S.N.) and two videos with immersive sound (S.I.), all with identical duration.

3.2. Experimental design and procedure.

Combining both forms of film post-production, image, and sound, we generated a total of four videos, with identical duration. Table 1 includes the detailed contents of videos.

Video 1: non-photorealistic - narrative sound (dialogues)
Video 2: non-photorealistic - immersive sound (ambients)
Video 3: photorealistic - narrative sound (dialogues)
Video 4: photorealistic - immersive sound (ambients)

Table 1. Image and sound post-production setups used in our tests.

These videos were presented to the participants, and their eye movements were measured with an eye tracker. Each volunteer watched the four videos using polarized glasses. The presentation sequence was varied to distribute the bias of seeing a film for a second, third, and fourth time. At the time of defining our experimental design, we did not have a previous hypothesis about how the gaze is affected simultaneously by the post-production of image and sound. If our study had focused only on the visual factor of NPR, we could have used a between-subjects design. However, despite the small size of our sample size, we used a within-subject design to open a possible future direction for experimentation and data gathering. At this stage of our research, our study does not intend to provide evidence extensive to all kinds of NPR styles that could be used in 3D cinema. In this sense, an expected continuation could be to study how visual artifacts can affect viewing patterns in the specific case of 3D cinema, with a larger number of NPR filters.

3.3. Data capture.

Each volunteer entered the laboratory, where they waited for a facilitator, who guided them during the entire process. The volunteers sat in front of the television and the eye tracker to watch the four videos in a varied order. The facilitator conducted a calibration phase with the eye tracker before presenting each video. The eye tracker generated for each (user, video) pair a text archive table with the columns timestamp, leftx, lefty, rightx, righty, pupilleft, pupilright. For this study, only the data for ocular focus was considered, leaving the data for pupil response for a future work.

3.4. Data processing.

Once the data capture process was completed we proceeded to unify those records into a file useful for posterior analysis. The first step in transforming this data was to annotate the files generated by the eye tracker with a user identifier and a video identifier. After that we could merge all files into a single one. Then, we transformed the timestamp column into the corresponding frame of the video. To do that, we split the video frame by frame, using the FFmpeg utility (www.ffmpeg.org/). Afterward, we filtered missing data and errors (data outside the monitor), and computed the average point of interest of both eyes. Finally, in order to increase the number of observations per frame, the showings of frame i was included in frame i-1 and i+1. The resulting dataset, containing identifiers for the user, video, frame, and the coordinates of the point of interest (x and y). It was composed for 1,349,944 validated records with an average of 62.14 captures per frame.
In order to check if what the users watched was significantly different, we performed a generalization of Student's t-statistic used in multivariate hypothesis testing called Hotelling statistic test [25] present in R packet ICSNP, version 1.1-0. To do so, for each pair of videos (V1-V2, V1-

V3, ..., V3-V4) and every frame, we tested the null hypothesis: what users are watching in video x and video y is the same. We also measured if the post-production produces more dispersion or diversity in what the users watched. We use the Shannon entropy [26], defined by an equation, as a measure of dispersion (see Table 2).

Before computing those values, we converted our two-dimensional variable, the coordinates of the point of interest, into a one -dimensional discrete variable. Both, entropy computing and variable transformation were calculated using the R entropy package, version 1.2.1.

$$\mathrm{H}(X) = \sum_{i=1}^{n} \mathrm{P}(x_i)\mathrm{I}(x_i) = -\sum_{i=1}^{n} \mathrm{P}(x_i)\log_b \mathrm{P}(x_i).$$

Table 2. Entropy defined as a measure of dispersion.

3.5. Data analysis.

In order to study the ocular response to audiovisual stimuli we combine quantitative data produced by the eye-tracker, with a qualitative analysis of the audiovisual content. The quantitative description was made in terms of two types of indicators: visual entropy and pupillary reactivity. The analysis of visual entropy in the 4 videos was made in temporal terms, considering the global values and the standard deviation for each video. Regarding the comparative analysis of the pupillary reactivity data in the 4 videos allowed the identification of three sets of results. We observed significant differences in pupillary reactivity between PR and NPR videos. We analyze the data to establish if there are possible correlations between the pupillary reactivity, the behavior of ocular fixations -described with the concept of entropy as a measure of attentional dispersion- and the temporal evolution of the film.

4. RESULTS

4.1. Pupillary reactivity to sound post-production.

Comparisons of the distributions were established to study the impact of the sound post-production on visual perception. No significant differences were observed that allowed to establish sound as a relevant factor in ocular response.

4.1. Comparison of NPR and PR videos in terms of visual entropy.

We established comparisons between the videos regarding their entropy (see Table 3), in order to identify if possible correlations exist between the visual style (NPR, PR) and the entropy of ocular fixations on the screen. We utilize a general comparison of the four videos regarding attentional entropy. The NPR videos tend to present lower levels of entropy than the PR videos. The result of this global comparison between videos is that NPR introduces a transversal difference, associated apparently with lower levels of entropy. See Figure 2.

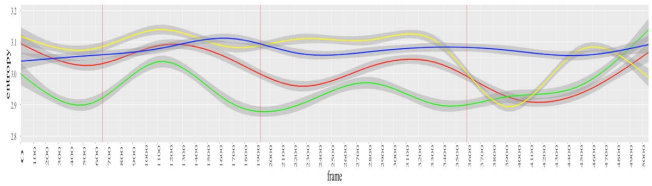

Figure 2. Attentional entropy: ocular response to the 4 short films in terms of visual entropy. NPR-IS [green curve], NPR-NS [red curve], PR-NS [blue curve] and PR-IS [yellow curve].

4.2. Pupillary reactivity to film post-production.

The pupillary reactivity was analyzed in front of the film sequence, eliminating those data associated with blinking and looking out of the screen, being filtered in is 0.01 πrad / samples. Video 1 (blue) and video 2 (red) correspond to the response of the presentation of the NPR video, while video 3 (cyan) and video 4 (magenta) correspond to the photorealistic presentation of the short film. See Figure 3.

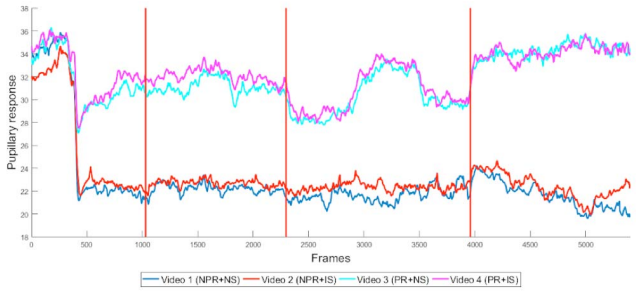

Figure 3. Average of the pupil response of 27 subjects compared to the presentation of the 4 short films. NPR-NS [dark blue curve], NPR-IS [red curve], PR-NS [blue curve] and PR-IS [pink curve].

4.3. Pupillary reactivity to NPR post-production.

Average of pupillary reactivity to sound design in PR and PPR videos. The envelope to the average corresponds to the 95% confidence interval obtained with a student test per frame. See Figure 4.

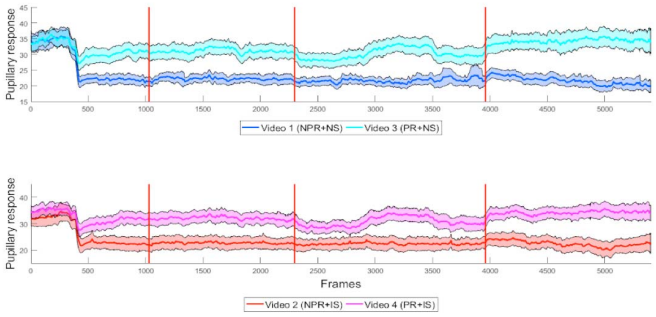

Figure 4. (A) Average of the pupil response for the presentation of the short film with sound type 1 in both NPR and PR. (B) Average of the pupil response for the presentation of the short film with type 2 sound in both NPR and PR. The envelope to the average has the same characteristics as in A)

4.5. Variance of light and pupillary reactivity to type of post-production.

We compared light variance in NPR vs PR, with these results: light NPR = 34,816 vs. light PR = 277,414. We also analyzed the consistency between variance of pupillary response according to post-production type. These are the values of light variance in the 4 videos.

Video 1 = 0,671 (NPR + NS) / Video 2 = 0.471 (NPR + IS)
Video 3 = 4,364 (PR + NS) / Video 4 = 3,494 (PR + IS).

4.4. Pupillary reactivity to the intensity of light and post-production.

The pupil average is contrasted with the intensity of light of the short films measured as the average of the levels of grays per frame, where the maximum possible value is 255 and the minimum 0. It is appreciated that on average the light level for the presentation PR (blue dotted line; mean = 71.3, sd = 16.6) is less than one third that of the NPR presentation (brown dotted line; mean = 223.2, sd = 5.9). See Figure 5.

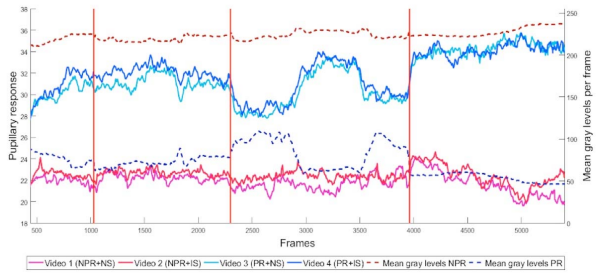

Figure 5. Pupillary reactivity to the intensity of light and post-production in NPR and PR videos.

5. DISCUSSION

Our previous research [3] enabled us to verify that, at least in the studied cases, the photorealistic style implies a greater dispersion of the gaze in the scene. At the same time, in the non-photorealistic mode we find users mainly focusing on faces and physical contact: we interpret this as a way to build a common thread of what is seen, from the perspective of each user. This is consistent with the idea of an audiovisual culture predominantly determined by the aesthetics of classical narrative cinema, where following a story constitutes the main intellective criterion. However, as we have already established in our previous work [3], this trend is not a rule, because it depends on the types of viewers. In a previous research on the perceptual and cognitive processes observed in Internet browsing [27] inferred from two intentional samples of 540 cases, one referring to young people and the other to adults, differentiated by social stratum, gender and ages of life, we detected three types of subjects based on their cognitive abilities: (a) a subject-type with cognitive simplicity able to do only one thing at a time, (b) a subject-type with a medium cognitive complexity able to work in four simultaneous spaces, and, (c) a subject-type with a high cognitive complexity capable of doing ten activities at the same time. Although the research referred to Internet browsing has samples collected from young people and adults at different times [27], they are a good source of hypothesis, especially in the following: it makes clear that young people tend to have medium and high cognitive complexity, with respect to cognitive simplicity observed in adults aged 40-45 and 50-55 years. This research also makes it clear that the perceptive-cognitive processes are in statistical correlation with the ages of life and not with the social stratum or gender. This element does not constitute in itself a specific interpretive criterion for our data since our users belong to only one age of life (18-23). However, it establishes the idea that there are different profiles of cognitive complexity within the same age group.

Considering these elements, the analysis of the pupillary reactivity observed in its temporality throws a concrete difference between photorealistic (PR) and non-photorealistic videos (NPR). In comparison with significative pupillary fluctuations observed in PR videos, the NPR filter used seems to produce a normalization effect pupillary reactivity to light fluctuations. This is confirmed by the light variance, which is much lower in NPR videos compared to PR videos. This tendency led us to expect that the pupillary reactivity would be determined by this lower light variance. Thus we expected a normalization effect of pupillary reactivity in PR videos, versus NPR videos where we observed that the behavioral patterns of pupillary reactivity differ from the light fluctuations. This led us to think that in NPR videos the pupillary reactivity does not seem to indicate an only response to light stimuli. Considering the acquired knowledge about cognitive complexity, we observe two relevant elements: (1) the

variance of pupillary reactivity is considerably higher in PR videos; however, it is suggestive that (2) in all videos, those videos that are associated with a narrative sound show a higher variance in pupillary reactivity. Although the sound was not significant as a relevant factor, the fact that the narrative sound who emphasized the voice of the characters, seem to have an effect in emotional response. This makes us think of a multimodal cognitive action that constructs an interpretative strategy of narrative totally based on the characters. This is consistent with the attentional evidence that points in the same direction: in the NPR videos, the characters concentrate more visual attention.

Finally, we find a result that we interpret from the perspective of cognitive complexity. In the final scene of the film, we detected the largest area of differences in pupillary reactivity in the NPR videos: the whole scene is built around a character who performs a doubly significant action: at the same time that she develops a verbal discourse more and more expressive in his facial gesture, she approaches the camera. During this body movement towards the camera, the stereoscopic effect produces the illusion that his head "emerges from the screen". The ocular task of constructing the stereoscopic illusion of an emergent face is added to the intellective action of constructing an interpretation of the scene based on facial and vocal expressivity. We know that this final scene is more visually complex than the previous ones. We also know that despite the subtraction of visual objects produced by the NPR filter, the pupillary behavior does not seem to be reacting only to light fluctuations. Considering the limited scope of our data, we cannot affirm that these differences correspond only to an emotional response, but it is clear that they do not seem to indicate a unique type of immersion. The emotional response seems to vary depending on the type of visual style, although the evidence is not so strong in this direction, it is clear that the research carried out opens implications beyond 2D and 3D cinema: it helps to understand cognitive complexity as a central factor to understand immersion. In this case, the greater ocular dispersion observed in the PR videos are consistent with a pupillary behavior globally more reactive to the light change. However, the exception found in the final scene, where NPR videos present a greater pupillary response, could indicate a feature consistent with the evidence mentioned above: greater visual complexity could be correlated with greater pupillary response. As the visual abstraction of the NPR implies less visual elements, we will understand the greater visual complexity as the accumulation of a sound element (rising voice), a visual element (facial expressions) and an optical element (emergent face in 3D). The idea that greater cognitive complexity can produce attentional and emotional reactions according to the type of users is consistent with previous results [12]: it was found that when in a 3D cinema scene a verbal story coincides with the stereoscopic illusion of an emerging image, the narrative comprehension of the scene varies greatly depending on the different cognitive complexity profiles of the users.

Then, everything is played between the space of the perceiving subjects and the type of visual style proposed. Our aim has been to describe with scientific basis how this interrelation between film and perception -immersion-, could be detected with an indicator that gathers attentional behavior and pupillary response. The gaze fluctuations (more entropic in PR videos than in NPR videos) may correspond with the corporal involvement indicated by the pupillary reactivity. Also, it could be inferred that when information is more dispersed, the intellective task of linking the sound treatment with the visual interpretation of the scene does not occur equally in PR and NPR videos. This led us to make the difference between a more emotional immersion, in NPR videos, than a more intellective immersion in PR videos.

The existing bibliography has addressed immersion from a phenomenological perspective through particular cases, which has led to inaccuracies or theoretical speculations. It is for this reason that our approach seeks to build a concept of immersion based on experimental evidence. Our objective is to produce an empirical measurement of immersion that allows us to give feedback to the filmmaker, regarding the ocular behavior triggered by a visual style, and in this way, to better enhance the emotional involvement of the spectators. By having an indicator to describe the ocular correlate of the immersion, we could reflect early on the relationship between post-production and emotion. Today at least we know that a non-photorealistic visual style could modify the behavior of our gaze, and with that, lead us to experience cinematic emotions.

6. CONCLUSIONS

From the preceding analysis, we can infer the verification of the initial hypothesis that it is possible to build an ocular indicator of immersion, which uses pupillary response information as a way of tracking the physical and emotional involvement of the viewer in the movie. The photorealistic visual style, which ocular response indicates greater attentional entropy, would have a more invariant development in its pupillary behavior: everything can be visually significant in the scene but the pupillary (emotional) response is smaller. It is a type of "analytical" immersion that seems to go through the scenic space in depth, observing the movement of the characters, but always assigning attentional resources to the visual exploration of the scene. At the same time, the inverse case is detected. The non-photorealistic visual style, whose attentional response indicates an attentional focus on the characters, presents lower intellective entropy. It is a type of "emotional" immersion where a greater visual attention seems to correlate to the scene interpretation with a pupil reactivity that suggests the emergence of an emotional response. Both

types of immersion suggest different modalities of corporal and ocular involvement during the temporal experience of the film.

7. FUTURE WORK

More studies are needed to be able to validate the applicability of this ocular indicator of immersion with other types of contents in stereoscopic 3D format.

8. REFERENCES

1. V.T. Visch, E.S. Tan, and D. Molenaar. The emotional and cognitive effect of immersion in film viewing, in Cognition & Emotion, 24: 8, pages 1439 — 1445, 2010.
2. V. Fajnzylber, M. Magdics, M. Castillo, C. Ortega, and M. Sbert. From 2D to 3D: A Case Study of NPR and Stereoscopic Cinema. In Smart Graphics. SG 2015. Lecture Notes in Computer Science, vol 9317, pages 87–98, 2017.
3. V. Fajnzylber, L. González, P. Maldonado, R. Del Villar, R. Yáñez, S.Madariaga, M. Magdics, and M. Sbert. Augmented film narrative by use of non-photorealistic rendering" In 2017 International Conference on 3D Immersion (IC3D), pages, 1-8. IEEE, 2017.
4. A.G. Patil, S. Raman. Automatic content-aware Non-Photorealistic Rendering of Images. In arXiv:1604.01962v4, pages 1-6, 2016.
5. M. Haller, C. Hanl, and J. Diephuis, J. Non-photorealistic rendering techniques for motion in computer games in Computers in Entertainment (CIE), 2(4), 1-10, 2004.
6. R.R. Henderson, M.M. Bradley and P.J. Lang. Emotional Imagery and Pupil Diameter. In Psychophysiology, 5(6), pages 1-7, 2018.
7. M. Bradley, L. Miccoli, Ma.A. Escrig, and P.J. Lang, P. J. The pupil as a measure of emotional arousal and autonomic activation. In Psychophysiology, 45(4), 2008.
8. A. Minassian, E. Granholm, S. Verney, W. Perry. Pupillary dilation to simple vs. complex tasks and its relationship to thought disturbance in schizophrenia patients. In International Journal of Psychophysiology, 52(1), pages 53-62, 2004.
9. Comparison of NPR and PR styles. Sample video 1. https://vimeo.com/234435709, 2017.
10. D. Mould, R. L. Mandryk, and H. Li. Emotional response and visual attention to nonphotorealistic images. Computers & Graphics, 36(6):658–672, 2012.
11. R. Sayeed and T. Howard. State of the Art Non-Photorealistic Rendering (NPR) Techniques. In Louise M. Lever and Mary McDerby, editors, Theory and Practice of Computer Graphics, The Eurographics Association, 2006.
12. T. Umenhoffer, L. Szecsi, and L. Szirmay-Kalos. Hatching for motion picture production. In Computer Graphics Forum, volume 30, pages 533–542. Wiley Online Library, 2011.
13. C. Richardt, J. E. Kyprianidis, and N. Dodgson. Stereo coherence in watercolour rendering. In Symposium on Non-Photorealistic Rendering and Animation 2010, June 2010.
14. R.A. Hamzah and H. Ibrahim. Literature survey on stereo vision disparity map algorithms. Journal of Sensors, 2016, 2015.
15. J. E. Kyprianidis and J. Döllner. Image abstraction by structure adaptive filtering. In Proc. EG UK Theory and Practice of Computer Graphics, pages 51–58, 2008.
16. H. Winnemöller, S.C. Olsen, and B. Gooch. Real-time video abstraction. ACM Trans. Graph., 25(3):1221–1226, jul 2006.
17. N. Redmond and J. Dingliana. A hybrid approach to real-time abstraction. In Proceedings of Eurographics Ireland, 2009.
18. A. Santella and D. DeCarlo. Visual interest and npr: An evaluation and manifesto. In Proceedings of the 3rd International Symposium on Non-photorealistic Animation and Rendering, NPAR '04, pages 71–150, New York, NY, USA, 2004. ACM.
19. J. E. Kyprianidis, J. Collomosse, T. Wang, and T. Isenberg. State of the 'art': A taxonomy of artistic stylization techniques for images and video. In Transactions on Visualization and Computer Graphics, 19(5):866–885, IEEE, May 2013.
20. Y. Kim, Y. Lee, H. Kang, and S. Lee. Stereoscopic 3d line drawing. ACM Trans. Graph., 32(4):57:1–57:13, July 2013.
21. L. Northam, P. Asente, and C.S. Kaplan. Stereoscopic 3d image stylization. Comput. Graph., 37(5):389–402, August 2013.
22. H. Winnemöller. Xdog: advanced image stylization with extended difference-of-gaussians. In John P. Collomosse, Paul Asente, and Stephen N. Spencer, editors, NPAR, pages 147–156. ACM, 2011.
23. M.Magdics, C. Sauvaget, R. Garcia, and M. Sbert. Post-processing NPR effects for video games. In 12th ACM International Conference on Virtual Reality Continuum: VRCAI 2013, pages 147–156, 2013.
24. Comparison of eye tracking results. Sample video 2. https://vimeo.com/234386284 , 2017.
25. H. Hotelling. A generalized T test and measure of multivariate dispersion. In Proceedings of the second Berkeley symposium on mathematical statistics and probability. The Regents of the University of California, 1951.
26. J. Lin. Divergence measures based on the Shannon entropy. IEEE Transactions on Information theory 37, no. 1, 145-151, 1991.
27. R. Del Villar. Proyecto Fondecyt No 1120064. Technical report, Consejo Nacional de Investigación