AUGMENTED FILM NARRATIVE BY USE ON NON-PHOTOREALISTIC RENDERING: FROM 3D CINEMA TO VIRTUAL REALITY

Victor Fajnzylber Reyes

DOCTORAL THESIS

AUGMENTED FILM NARRATIVE BY USE OF NON-PHOTOREALISTIC RENDERING: FROM 3D CINEMA TO VIRTUAL REALITY

Victor Fajnzylber Reyes

2020

University of Girona

Doctoral Programme in Technology

AUGMENTED FILM NARRATIVE BY USE OF NON-PHOTOREALISTIC RENDERING: FROM 3D CINEMA TO VIRTUAL REALITY

DOCTORAL THESIS

Report presented by Victor Fajnzylber Reyes in order to obtain

the degree of Doctor by the University of Girona

Director: Mateu Sbert Casasayas

2020

Universitat de Girona

CERTIFICATE

Girona, May 3 th, 2021

Hereby, I, Mateu Sbert, Director of the GILAB, Graphics and Imaging Laboratory, a research group in computer graphics at the University of Girona, extend the following certificate: the doctoral research carried out by Victor Fajnzylber Reyes, under my direction, has come to an end, with the presentation of this thesis to qualify for the degree of Doctor of Technology at the University of Girona, Spain. Currently, Mr. Fajnzylber is in the third year of his doctorate and has finished writing his doctoral thesis, in English, under the title "Augmented film narrative by use of non-photorealistic rendering: from 3D cinema to virtual reality".

This thesis has an interdisciplinary and international character: different collaborations have produced scientific research about cinematic immersion and virtual reality, in the crossroads of computer graphics, neuroscience, sociology and film studies. Mr. Fajnzylber's research has resulted in the publication of five peer reviewed articles, three presentations of articles at international conferences, the organisation of an international conference on the use of eye-tracking in immersive technologies, and carrying out research stays in virtual reality research laboratories in Chile, France and Belgium. These activities have demonstrated that Mr. Fajnzylber possesses the necessary competencies to conduct empirical and experimental research on technology.

This certificate is extended to be included in the doctoral thesis.

Dr. Mateu Sbert Casasayas

Professor of Computer Languages and Systems
Director of the Graphics and Image Laboratory
Institute of Informatics and Applications, University of Girona
PIV Building, Montilivi Campus, 17071, Girona / Tel: +34 972 41 84 19

To my wife Karina

To my sons Diego and Ruy

To my brothers Pablo and Guayo

Victor Fajnzylber Reyes

May 2021

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Foreword

The first inspiration for this thesis was a conversation with Professor Mateu Sbert in Santiago, the Chilean capital, some years ago. We discussed perception and computer graphics. We wondered if a computing technology such as non-photorealistic rendering (NPR) could modify the cognition of a visual based narrative such as cinema, video games and virtual reality. The question at that time was: could NPR generate a "new experience" of the same narrative? That was the beginning of this doctoral research, conceived as interdisciplinary, in the crossroads between computer science, film studies, sociology and neuroscience. Years later, finally, the original hypothesis seems to be confirmed with experimental evidence: the NPR is correlated with a specific ocular behavior, focused on characters and their body interactions, enhancing the narrative dimension of the cinematic experience. Non-photorealistic rendering modifies the visual perception of stereoscopic scenes, and thus, seems to trigger a new experience of immersion of the same cinematic narrative. This is the main idea of this doctoral thesis.

I want to thank my supervisor Mateu Sbert for his guidance during these years. I am grateful for having the opportunity to confirm a personal belief: interdisciplinary research appears to me as the best way to explore and learn new forms of creativity. In this thesis we combined scientific methods of computer graphics and neurosciences with artistic experimentation in 3D cinema and virtual reality, to study the immersive effects of non-photorealistic rendering (NPR). We have structured an interdisciplinary approach about NPR and I hope this work could be useful in research fields related with visual perception and computer graphics.

Based on the direction of our future work, I think that the main applications of this thesis could be oriented to applied research and production of narrative video games, interactive virtual reality simulations and cinematic immersive experiences.

During the time it took to complete this thesis, the following activities were carried out, in order to enrich the point of view of cinematic perception and virtual immersion.

I have spent research stays on virtual reality in the following laboratories, centers, events and projects.

The contribution of these activities was to improve the understanding of virtual reality, in its dimensions of perception, cognition, production and reception by users.

1) Centre de Réalité Virtuelle de la Méditerranée / CNRS (Marseille, France, 2018).
2) Centre Européen de Réalité Virtuelle / Lab-STICC (Brest, France, 2018).
3) HYBRID INRIA, immersive technologies lab, (Rennes, France, 2018).
4) IMMERSIA, immersive technology platform (Rennes, France, 2018).
5) VR DAYS, immersive technologies event in Amsterdam (Holland, 2018).
6) STEREOPSIA, immersive technologies event in Brussels (Belgium, 2018).
7) SPATIAL MEDIA LAB, immersive technologies lab, Ecole Nationale Supérieure d´Arts Décoratifs (Paris France, 2019).
8) LUTIN USERLAB, Cognition Laboratory, Cité des Sciences et de l´Industrie (Paris, France, 2019).
9) Université de Liège, Programme Teaching with VR (Liège, Belgium, 2019).
10) Cabinet de Réalité Virtuelle, Museúm National d´Histoire Naturelle (Paris, France, 2019).
11) Département Numérique, Musée D´Orsay (Paris, France, 2019).
12) Institut Français, Focus en Médiation Numérique (Paris, France, 2019).
13) Festival PIXII, Sunny Side of the Doc (La Rochelle, France, 2019).
14) International France-UK Forum on Virtual Reality and Museums (Paris, France, 2019).
15) Festival New Images, Forum des Images (Paris, France, 2019).

We also presented papers in international conferences. The contribution of these activities was to improve the understanding of my research theme in the context of state of the art studies on immersive languages and technologies.

1) The Smart Graphics Conference 2015, Chengdu, China (paper presentation).

2) The International Conference on 3D Immersion 2017, Brussels, Belgium (paper presentation).

3) The International Conference on 3D Imaging 2018, Brussels, Belgium (paper presentation).

4) The International Conference on 3D Imaging 2018, Brussels, Belgium (thematic conference). In addition to the presentation of articles, I was invited to act as "Chair" of the thematic conference "Eye tracking and Analysis for XR"[1] , where we brought together researchers from Belgium, United Kingdom, Chile, Luxembourg and Spain, to discuss the uses of eye-tracking to study perception in virtual reality and in immersive technologies.

[1] (XR = Extended Realities)

Publications. The contribution of the following publications to this thesis was to systematize, analyze and present the results obtained experimentally during the period of this research (2017-2020).

1) Title: *"From 2D to 3D: A Case Study of NPR and Stereoscopic Cinema"*, published in "LECTURE NOTES ON COMPUTER SCIENCE", Springer (March 2017). Authors: Víctor Fajnzylber, Milán Magdics, Macarena Castillo, Constanza Ortega and Mateu Sbert.
Link: https://link.springer.com/chapter/10.1007%2F978-3-319-53838-9_7

2) Title: *"Augmented film narrative by use of non-photorealistic rendering"*, published in "IEEE Xplore Digital Library" (March 2018), 2017 International Conference on 3D Immersion, IC3D - Proceedings. Authors: Víctor Fajnzylber, Larry González, Pedro Maldonado, Rafael del Villar, Rodrigo Yáñez, Samuel Madariaga, Milán Magdics and Mateu Sbert.
Link: https://doi.org/10.1109/IC3D.2017.8251912

3) Title: *"Pupillary reactivity to non-photorealistic rendering: a case study of immersion in 3D cinema"*, published in "IEEE Xplore Digital Library" (March 2019), 2018 International Conference on 3D Immersion, IC3D - Proceedings. Authors: Víctor Fajnzylber, Samuel Madariaga, Rafael del Villar, Pedro Maldonado, Diego Vargas, Aria Serra, Milán Magdics and Mateu Sbert. Link: https://doi.org/10.1109/IC3D.2018.8657894

4) Title: *"From NPR to VR: tracking ocular behavior in immersive virtual reality"*, published in "Communication Papers" journal, University of Girona (December 2019). Authors: Víctor Fajnzylber, Andrea Wenner, Javier Moyano and Mateu Sbert.
Link: https://communicationpapers.revistes.udg.edu/communication-papers/article/view/22328

5) Title: *"Transdisciplinary study of immersion in interactive virtual reality"* was submitted and accepted for publication in the french journal "Revue Française de Méthodes Visuelles", n°5, "Images interactives et nouvelles écritures" (accepted in September 2020 for publication in 2021). Authors: Víctor Fajnzylber, Francisco Gutiérrez, Paulo Barraza, Pablo Riveros, Javier Moyano and Mateu Sbert. Link: https://rfmv.fr

All these activities became contributions to this interdisciplinary thesis.

Acknowledgements

I am very thankful to my supervisor Mateu Sbert, for his support and encouragement, through the development of this thesis.

I want to thank the support of Abril Alzaga and Gabriel Rodriguez of the Ingmar Bergman Cinema and Theater Chair of the National Autonomous, University of Mexico (UNAM) and Alexandre Saudinos of Parrallell Cinéma (France) for their help in the shooting of a stereoscopic film in the house of Luis Buñuel in Mexico City. This film was an important asset for the experiments with eye-tracking.

I want to thank all the people with whom I was fortunate to collaborate, learn, create images, design interactions, and write together in the five publications generated throughout this interdisciplinary doctoral research. Thanks to you I had the opportunity to create an interdisciplinary thesis, combining perspectives, concepts and methods from disciplines such as film, computer graphics, sociology and neurosciences. All my gratitude to Milán Madgics, Rafael del Villar, Larry González, Pedro Maldonado, Aria Serra, Rodrigo Yáñez, Samuel Madariaga, Francisco Gutiérrez, Paulo Barraza, Pablo Riveros, Macarena Castillo, Contanza Ortega, Diego Vargas, Cristobal Maldonado and Daniel Collao, who actively participated in the creation of images and interactions of the prototype of the virtual reality, and many thanks to Javier Moyano and Andrea Wenner for their precious assistance.

Special thanks to Tim Smith, of the Birkbeck University of London for his precious advices about eye-tracking, David Grogna of the Université de Liège and Karina Sequeira of the University of Chile, for their scientific advice, and Clara Szwarc of the International Relations Unit of the Ministry of Cultures, Arts and Heritage of Chile for his support to assist to the scientific conference of Stereopsia 2017.

The research for this thesis was in part sponsored by grants TIN2016-75866-C3-3-R and PID2019-106426RB-C31 from the Spanish Ministry of Economy and Competitiveness, 2014SGR1232 from Catalan Government, from the Direction of Research of the Institute of Communication and Image, and the support of Research Direction of the University of Chile.

Abstract (*English*). It is stated that non-photorealistic rendering (NPR) triggers specific ocular behaviors that can have a real effect on the way of perceiving a cinematic narrative. The transdisciplinary interest of this thesis lies in the fact that the concept of "film narrative" is not something exclusive to cinema, but that can also be conceived as a set of audiovisual procedures applicable to virtual reality and video games. The fundamental idea is that NPR post-production provides a visual style that can modify the way of perceiving a narrative, beyond the audiovisual format where it is applied. The empirical research process is made up of three stages: (1) we questioned about the possible influence of NPR on cinematic perception in 2D cinema and stereoscopic 3D photography; then (2) we were interested in the ocular behavior, attentional and pupillary, of 27 viewers faced to a film produced in stereoscopic 3D format with NPR post-production, and finally, (3) we projected these results to study their application in the production of stereoscopic 3D virtual reality. The combination of NPR and stereoscopy is not a perceptual contradiction. Our experimental results in 2D cinema indicate that narrative comprehension is not significantly affected by NPR. In order to verify this first result, which points to the effect of perceptual modifications produced by NPR, we elaborated an experimental design based on the use of eye-tracking to collect data about ocular behavior of spectators of a 3D film. This allowed us to analyze ocular behavior in two dimensions: attentional and pupillary reactivity to NPR. We concluded that a NPR filter associated with a high level of abstraction could trigger an adaptive ocular response that may be described as perceptual anxiety: a lower exploration of the scene and a greater focus on the more expressive elements of the scene: bodies and faces of the characters. This ocular behavior, triggered by NPR, could benefit cognitive immersion, especially in cinematographic genres (fantasy, horror, suspense, science fiction) that use visual abstraction and cognitive uncertainty as an aesthetic effect to favor greater involvement of viewers in their narrative movie. The contribution of our study of visual perception lies in the understanding of the user-viewer: we propose a way of interpreting the relationship between visual style and immersion. Our results were applied to virtual reality production, in order to produce an immersive content, were the user behavior could be studied, using eye-tracking, as a way to guide the search for a better immersion in the proposed narrative. We combine scientific research and audiovisual production in a way that contributes to the creation of new knowledge about visual perception in 3D cinema and virtual reality.

Resumen (*Castellano*). Se afirma que el renderizado no-fotorealista (NPR) desencadena comportamientos oculares específicos que pueden tener un efecto real en la forma de percibir una narrativa cinematográfica. El interés transdisciplinar de esta tesis radica en que el concepto de "narrativa cinematográfica" no es algo exclusivo del cine, sino que también puede concebirse como un conjunto de procedimientos audiovisuales aplicables a la realidad virtual y los videojuegos. La idea fundamental es que el NPR aporta un estilo visual que puede modificar la forma de percibir una narrativa, más allá del formato audiovisual donde se aplica. Nuestra investigación empírica consta de tres etapas: (1) evaluamos la posible influencia del NPR en la percepción cinematográfica en el cine 2D y la fotografía 3D estereoscópica; luego (2) nos interesamos por el comportamiento ocular, atencional y pupilar de espectadores de una película producida en formato 3D estereoscópico con postproducción NPR, y finalmente (3) proyectamos estos resultados para estudiar su aplicación en la producción de realidad virtual 3D estereoscópica. La combinación de NPR y estereoscopía no es una contradicción perceptiva. Nuestros resultados experimentales en cine 2D indican que el NPR no afecta significativamente la comprensión narrativa. Para verificar este primer resultado elaboramos un diseño experimental basado en el uso del eye-tracking para recolectar datos sobre el comportamiento ocular de los espectadores de una película en 3D. Esto nos permitió analizar el comportamiento ocular en dos dimensiones: la reactividad atencional y pupilar al NPR. Concluimos que un filtro NPR asociado a un alto nivel de abstracción podría desencadenar una respuesta ocular adaptativa que puede describirse como "ansiedad perceptual": una menor exploración de la escena y un mayor enfoque en los elementos más expresivos de la escena, cuerpos y rostros de los caracteres. Este comportamiento ocular, desencadenado por NPR, podría beneficiar la inmersión cognitiva, especialmente en géneros cinematográficos que utilizan la abstracción visual y la incertidumbre cognitiva como efecto estético (fantasía, terror, suspenso, ciencia ficción) para favorecer una mayor implicación de los espectadores en su narrativa. La contribución de nuestro estudio de la percepción visual radica en la comprensión del usuario-espectador: proponemos una forma de interpretar la relación entre estilo visual e inmersión. Nuestros resultados se aplicaron a la producción de realidad virtual, con el fin de producir un contenido inmersivo, donde se pudiera estudiar el comportamiento del usuario, utilizando el seguimiento ocular, como una forma de orientar la búsqueda de una mejor inmersión en la narrativa propuesta. Combinamos la investigación científica y la producción audiovisual de manera que contribuya a la creación de nuevos conocimientos sobre la percepción visual en el cine 3D y la realidad virtual.

Resum (*Català*). S'afirma que el renderitzat no-fotorealista (NPR) desencadena comportaments oculars específics que poden tenir un efecte real en la forma de percebre una narrativa cinematogràfica. L'interès transdisciplinar d'aquesta tesi radica que el concepte de "narrativa cinematogràfica" no és una cosa exclusiva de cinema, sinó que també es pot concebre com un conjunt de procediments audiovisuals aplicables a la realitat virtual i els videojocs. La idea fonamental és que el NPR aporta un estil visual que pot modificar la forma de percebre una narrativa, més enllà de el format audiovisual on s'aplica. La nostra investigació empírica consta de tres etapes: (1) avaluar la possible influència de l'NPR en la percepció cinematogràfica al cinema 2D i la fotografia 3D estereoscòpica; després (2) ens interessem pel comportament ocular, atencional i pupil·lar d'espectadors d'una pel·lícula produïda en format 3D estereoscòpic amb postproducció NPR, i finalment (3) projectem aquests resultats per estudiar la seva aplicació en la producció de realitat virtual 3D estereoscòpica. Suggerim que la combinació de NPR i estereoscòpia no és una contradicció perceptiva. Els nostres resultats experimentals en cinema 2D indiquen que el NPR no afecta significativament la comprensió narrativa. Per verificar aquest primer resultat vam elaborar un disseny experimental basat en l'ús de l'eye-tracking per recol·lectar dades sobre el comportament ocular dels espectadors d'una pel·lícula en 3D. Això ens va permetre analitzar el comportament ocular en dues dimensions: la reactivitat atencional i pupil·lar al NPR. Concloem que un filtre NPR associat a un alt nivell d'abstracció podria desencadenar una resposta ocular adaptativa que pot descriure com "ansietat perceptual": una menor exploració de l'escena i un major enfocament en els elements més expressius de l'escena, cossos i rostres de els caràcters. Aquest comportament ocular, desencadenat per NPR, podria beneficiar la immersió cognitiva, especialment en gèneres cinematogràfics que utilitzen l'abstracció visual i la incertesa cognitiva com a efecte estètic (fantasia, terror, suspens, ciència ficció) per afavorir una major implicació dels espectadors en la seva narrativa. La contribució del nostre estudi de la percepció visual rau en la comprensió de l'usuari-espectador: proposem una forma d'interpretar la relació entre estil visual i immersió. Els nostres resultats es van aplicar a la producció de realitat virtual, amb la finalitat de produir un contingut immersiu, on es pogués estudiar el comportament de l'usuari, utilitzant l'eye-tracking, com una forma d'orientar la recerca d'una millor immersió en la narrativa proposta . Combinem la investigació científica i la producció audiovisual de manera que contribueixi a la creació de nous coneixements sobre la percepció visual en el cinema 3D i la realitat virtual.

In the 20th century, film narrative evolved enormously with the advent of digital post-production. New image processing and editing technologies allowed to multiply the expressive forms that could be obtained by modifying the visual flow generated during the shooting phase. In the case of editing, from analog procedures to digital operations, the cinema production process achieved a better integration of all phases. Digital editing brought greater efficiency but the language was not substantially modified, keeping the main function of producing the perceptual illusion of space-time continuity, in order to create the cinematic narrative, by stitching disconnected fragments. The viewer perceives a continuity of images and sounds for a certain time: we call that experience "watching a movie". Faced to a film narrative, the viewers assume a priori that the work is a continuous and coherent whole [Smi12]. To have this impression, viewers do not need to understand the technical way in which this illusion of cinematic continuity is built. At the same time, the inability to detect continuity errors shows that visual attention is not paid to all the details of the scenes. Viewers focus their attention on the details that seem relevant to them and according to the narrative proposed in each film. In other words, the viewer modulates their gaze behavior based on "objects of interest", guided by the narrative hierarchies that each film assigns to its characters, distinguishing the protagonists from the secondary characters. Various studies have shown that the viewer will focus his/her attention on the character's actions, and in particular on their faces, rather than on peripheral details [Smi12] [SM13].

A filmmaker can use cinematographic language, combining narrative events (actions of characters) and cinematic operations (such as montage or musicalization), to guide the viewer's gaze, seeking to generate moments of "attentional synchrony" between the viewers that make up an audience [Smi12]. Attentional synchrony could be defined as a temporal synchronization and spatial convergence of attentional behavior, of a significant part of an audience. Their gaze is convergent at the same time and in the same place of the image. However, this attentional synchrony is not a lasting nor a general condition of film perception, but rather cohabitates with its opposite behavior, that is, visual entropy, defined as dispersion of the gaze in the visual space of the film. Despite these advances in the study of film cognition, many questions remain about how we perceive a film.

One of the topics that still requires further investigation is the perception of the film narrative, influenced by the post-production decisions. The story is built during the edition, with the photographic and sound material during production, but it will be during the post-production phase where the narrative and expressive qualities of the story are enhanced. The same story, when presented in different visual styles, is modified through the impact of color, lighting, texture, and other visual attributes. Based on the emotional and cultural background of the viewer, the perception of these elements will allow to modify the cinematographic experience: each person could have a particular combination of feelings or recallings with colors or lighting.

It is a fact that in all film production, at the end of the edition, a process of trial and error begins, where different visual attributes are tested until defining the final aesthetics of the film. This is particularly true for 3D cinema productions, where editing and post-production must consider the ocular efforts that stereoscopy triggers on the viewer. The same film, when presented in 3D format, will tend to last longer: the human visual system needs more time to adapt to stereoscopic images than in 2D images. This is valid for different types of screens. Stereoscopic film productions will seek to obtain the greatest visual comfort, that is, that stereoscopy becomes a positive attribute that contributes to the aesthetic experience, doing everything possible to suppress discomfort caused by sudden changes in visual depth. With the exception of some experimental works, the vast majority of 3D cinema films are characterized by being presented with photorealistic visual styles. In historical terms, 3D cinema and photorealism have always existed together. However, in order to discover new possibilities of cinematic expression, we have been interested in exploring the effects that non-photorealistic post-production may have on the perception of film narrative, with special interest in stereoscopic cinematography.

In order to understand the effects of Non-Photorealistic Rendering (NPR) in film narrative perception, we will use the concept of “immersion”, widely used to analyze the perception and cognition of different artistic languages, from literature to cinema, video games and virtual reality (VR). For the purposes of our research, we are interested in the concept of immersion as it is used in the field of virtual reality. Our research on NPR perception is initially focused on stereoscopic cinema, as a way to infer possible insights for a better understanding of film narrative in stereoscopic virtual reality.

Within the scientific literature, the concept of “immersion” comes from the idea of experiencing physical engagement in a given environment. In a field familiar with this term, such as virtual reality, it is related to the perception of feeling physically “present” in a simulated environment. Virtual reality conveniently provides input devices, such as headsets and helmets, almost natural means of interaction and control, with a degree of perception of self-movement, which, in turn, allow the senses to be isolated enough to feel transported to another place. Good practices can increase this degree of sensation in the user experience. Immersion is studied in different fields and a virtual reality experience may not be immersive. For example, the sensation of immersion could be present in different experiences such as reading a book or a comic strip. However, for the purposes of our research, immersion can also be understood as the objective level of sensory fidelity that a virtual reality system provides [Sla03]. That is, the ability of a simulation to approach a real experience. For this, the sensations of the real world must be eliminated as much as possible, and replaced by the sensory experiences corresponding to virtual environments [Mes05]. “Presence”, often correlated and at the same time confused with “immersion”, is understood as the user's subjective psychological response to simulation [Sla03] [Ded09]. It is a cognitive and perceptual consequence of the immersion, and even understood as the last phase of it, after the stages of commitment and absorption of experience [BC04]. In simple terms, it corresponds to an illusion; the subjective belief that you are in one place, even when you are physically located in another [WS98] [PCSRVC00].

From the above, it follows that one cannot speak of immersion, and consequently, of “immersive” experiences without hinting at the sensory and cognitive impact, in short, the physical effect produced by these digital devices. Compared to the initial decades of these technologies, digital high definition, computational processing speeds, the ergonomics of new headsets and globalized access to these types of devices would suggest that the virtual era has arrived. Even the Covid-19 pandemic has been a booster for this VR technology. However, despite the enormous technological advance, problems of discomfort and various types of negative side effects from VR contents and simulations have still been reported in users. However, these problems have been extensively studied by the academy, generating metrics and standardized procedures to be able to assess these discomforts. And with this, improve experiences to minimize these effects, which mostly have to do with depth fields, experience control and color tone within the experience [BBI13] [SI20].

The inseparable character between virtuality and corporality has not been resolved, not even in industrialized countries where immersive technologies are more consolidated. This occurs in the artistic-cultural content (films, games, educational experiences) as well as in the experiences destined to the industrial training of labor competences. Virtual reality experiences still produce physical discomfort since all the keys to the associated physiological phenomena have not yet been deciphered, like in immersive flight simulators [KA19], much less in the field of content producers or l. In other words, creators and engineers produce content without possessing methods or tools that allow them to evaluate the physiological impacts of their products. Large sums of money can be spent to produce immersive experiences that, ultimately, produce side effects, without really knowing how to avoid them, except by going through trial and error processes that practically no independent audiovisual producer can afford [GB18] [KHP20] [ODBK18].

From all of the above, as a context, the importance of being able to develop a strategy to scientifically measure immersion is clear: developing a method to describe the involvement of the user's body during their experience with a 3D cinema or virtual reality film. This is the proposal of this research: we will present our step by step method to study immersion based on visual perception and ocular behavior. Thanks to this study of the user's perceptual and cognitive immersion in audiovisual content, we will be able to build the validity of our thesis: the use of non-photorealistic post-production produces a significant change in the way of experiencing filmic narrative. This perceptual change is manifested through the emergence of new patterns of ocular behavior, which indicate a greater focus on the narrative and the characters.

The objective of this thesis is to answer how NPR post-production style can modify the visual and cinematographic narrative of an immersive experience, and with it, contribute to an “increased” experience of its users, evaluable through the quality of the experience and the eye tracking. For this, it is preliminarily argued that: NPR can act as a factor of immersion, since it would add a “new layer” to the visual and film narrative. For this reason, its subordination to experiences in immersive fields of vision, such as in 3D technology and virtual reality, could favor the understanding of both the adjustments and the mismatches between formal innovations in audiovisual and multimedia matters, and cognitive reactivity and body response of viewers.

In this thesis we will present the foundations of the method of construction that combines both audiovisual production and experimental research, destined to generate knowledge about the individual immersive experience in an interactive virtual reality experience. The question on "how to understand the virtual" continues. We will not try to resolve it in this thesis but we will, at least, empirically outline some of the elements of the "immersive" phenomenon.

Doctoral research objectives

General objective

To assess the perceptual and cognitive response of viewers to film content processed with non-photorealistic image processing technology, through the study of correlations between ocular physiological indicators, film analysis and subjective self-evaluation reports of the experience.

Specific objectives

1. To implement an experimental measurement method to describe the ocular behavior of spectators during the cinematic viewing experience, through the use of eye-tracking techniques.

2. To develop a multimodal analysis method that correlates ocular behavior data of spectators and the temporal development of the film narrative in terms of its cinematographic events and procedures.

3 . To create a transdisciplinary approach for the study of the immersion of spectators in stereoscopic audiovisual content, which allows to transform the descriptions of ocular behavior into results applicable to the production of immersive content for an interactive virtual reality format.

Structure of this thesis

To build this NPR visual perception study, we have organized the research in the following structure of chapters.

Chapter 2 / Related research / Perceptual effects of non-photorealistic rendering.

In this chapter we present the previous research of interest as a reference to be able to build our object of study. In general terms, we will address the different dimensions related to NPR and filmic perception, which constitute the conceptual pillars of this thesis. We will approach NPR as a technique, visual style and particular modality of visual perception of the image. Also, the first relevant concepts related to filmic perception in stereoscopic 3D cinema and virtual reality will be presented. Some concepts will be taken up in later chapters, when defining the experiments and the empirical data generated to operationalize the relevant variables, such as immersion, visual attention, pupil reactivity, among others.

Chapter 3 / Methods / Using eye-tracking to study perception of NPR in 3D cinema.

In this chapter we present the methods used to empirically evaluate the research questions, combining cinematographic analysis with the observation of ocular behavior in real-time. In some cases, additional methods will be presented in greater detail in the results chapters, to contextualize, on a case-by-case basis, the way in which the data from this research was obtained.

Chapter 4 / Results / Perceptual effects of NPR: from 2D to 3D cinema

In this chapter we present the first results of the research: the starting point to study the perception of NPR in 2D and 3D cinema. Combining the NPR processing of film fragments and stereoscopic frames, with surveys, interviews and focus groups, it was possible to infer phenomena related to NPR and its effects on the way of perceiving the moving image.

The results of this chapter are published in Lecture Notes on Computer Science (March 2017), with the title "*From 2D to 3D: A Case Study of NPR and Stereoscopic Cinema*" (see here)

Chapter 5 / Results / Attentional reactivity to NPR in 3D cinema.

In the chapter we present experimental results associated with the attentional dimension of ocular behavior, thanks to data generated with eye-tracking technologies. We discuss how NPR modifies the way of capturing visual information while viewing a 3D movie. The main result is the inference of eye behaviors that may be of interest for understanding the effects of NPR on film perception. The results of this chapter are published in IEEE Xplore Digital Library (March 2018), with the title "*Augmented film narrative by use of non-photorealistic rendering*" (see here)

Chapter 6 / Results / Pupillary reactivity to NPR in 3D cinema.

In this chapter we present experimental results associated with the pupillary dimension of ocular behavior, thanks to data generated with eye-tracking technologies. We discuss how NPR modifies the way in which light information is captured when viewing a 3D movie. The main result is the inference of pupillary response behaviors that may be of interest for understanding the effects of NPR on film perception. The results of this chapter are published in IEEE Xplore Digital Library (March 2019), with the title "*Pupillary reactivity to non-photorealistic rendering: a case study of immersion in 3D cinema*" (see here)

Chapter 7 / Discussion / Using NPR to study visual perception in virtual reality.

To better understand the effects of NPR on film perception, in this chapter we present a reflection on the interest in conceptually integrating the attentional and pupillary dimensions of ocular behavior. We propose this integrated approach to NPR visual perception to build an operational, measurable concept of filmic immersion. We believe that the bodily manifestation of immersion, described from the observation of ocular behavior during the temporal evolution of the user experience, informs us about the hybrid effect, both attentional and emotional, of NPR on filmic perception. The main contribution of this chapter is to relate the ocular behavior as an indicator of immersion in the filmic narrative, identifying a potential application for cinema and virtual reality. The results of this chapter are published in Communication Papers, University of Girona (December 2019), with the title "*From NPR to VR: tracking ocular behavior in immersive virtual reality*" (see here)

Chapter 8 / Application / A transdisciplinary method to study immersion in virtual reality.

In this chapter we present an application of our results, about visual perception of NPR, for the implementation of a visual method to produce interactive virtual reality. By establishing as a premise that the visual style has effects on the way of seeing a scene, we could imagine, during the production process, different mechanisms of interaction, implicit or explicit, that could be based on expected ocular behavior. Considering that the difference between interactive virtual reality and videogames lies mainly in the narrative, and not so much in the interaction design, it can be inferred that the application of this visual method to virtual reality could be potentially useful for applications in the field of entertainment and educational video games. The results of this chapter are accepted as a paper in Revue Française de Méthodes Visuelles (2021), N°5, "Images interactives et nouvelles écritures", with the title "*Transdisciplinary study of immersion in interactive virtual reality*" (see here)

Chapter 9 / Conclusions / Augmented film narrative by use of NPR.

In this chapter we summarize the research findings. In general terms, it is stated that NPR triggers specific ocular behaviors that can have a real effect on the way of perceiving audiovisual work. The transdisciplinary interest of this thesis lies in the fact that the concept of "film narrative" is not something exclusive to cinema, but can be conceived as a set of audiovisual procedures applicable to virtual reality and video games. The fundamental idea is that image post-production provides a visual style that can modify the way of perceiving a narrative, beyond the audiovisual format where it is applied.

Chapter 10 / Future work / Using eye-tracking to study perception in virtual reality.

In this chapter we present the applied research goals that we will develop, as a scientific projection of this thesis, in the Virtual Realities Lab (VR-LAB), created in October 2020, at the University of Chile. The use of eye tracking to study the impact of NPR in 3D cinema has provided us with evidence of great interest to explore and describe the sensation of cinematic immersion in virtual reality.

The main research field in this VR-LAB is the combination of arts and sciences to apply this knowledge about ocular behavior to the production of narrative video games, interactive virtual reality simulations and cinematic immersive experiences.

2.1. NPR and Stereoscopy

Within graphics computing, "photorealistic rendering" (PR) has been a traditional area of image and stage production, which aims to simulate the optical attributes of a photographic camera. For the creation of this type of visual representation, it takes into account the understanding of the different factors that determine whether the viewer really perceives it as photographic [Rad02], increasing their sense of psychological presence in this representation [ZMM19]. However, there is another trend, within graphic computing, which groups together various algorithms, or visual representation techniques that go beyond this photorealistic paradigm [Agr09].

"Non-photorealistic rendering" (NPR), through its processing and filtering techniques, has been used, just like PR, for the production of images and scenes. This has allowed visual storytellers to expand into a variety of styles, focused on the abstraction and stylization of their content [KPNM19]. A visual abstraction corresponds to a transformation of the visual information data with some loss [VCI19].

Each of the NPR algorithms determines which informational aspects should be removed and which should be retained and made stylistically more prominent. The NPR set provides a wide range of styles, with aesthetic qualities such as drawing, illustration or painting [KPNM19]. These and more styles have been used for practical purposes and aesthetic reasons, from understanding their perceptual effectiveness in a virtual space [GW02] to creating an immersive cinema [CDLK19].

NPR conventionally has stylized, abstract representations in two-dimensional (2D) images and films. However, in recent times, processing and filtering techniques have been produced that can be used in the transformation in stereoscopic visualization [Sta08] [TCNN10]. This, in relation to traditional visual media, could facilitate the creation of more aesthetically elaborated works of art with greater involvement of the spatial aspect. In turn, it would allow more effective depth information to be delivered, collaborating with the identification and separation of depth layers in a scene, for example.

In this regard, stereoscopic 3D (S3D) is a technique capable of generating an illusion of depth in a 3D image, from two slightly different 2D images, one for each eye. This binocular addition technique, through visual devices such as anaglyphs, shutter glasses and head-mounted display, intensifies perceptual realism. This allows it to be used as a resource to offer an immersive feeling in three-dimensional virtual environments [SM15].

Regarding the combination between NPR and S3D, it is believed that this could improve depth perception, that is, the related distance between the visual object and the observer, and provide an improved visual experience for viewers [LB14]. The 3DS, therefore, would modify the perceived quality of the image, either by adding a feeling of "being there" [DB10].

2.2. Perception effects of NPR

Within the research line on the perception of PR scenes and artificiality, according to evidence [SN12], the visual system detects transformations from natural images to artificial images more easily than in its opposite process: it is more attentive to disappearing naturalness or emerging artificiality, especially in relation to faces. From another perspective, recognition of abstract images is significantly faster than with real photographs [GRG04] [WOG06]. In these and other cases, visual attention is directed, either to regions of space or to an object, selecting relevant information and filtering the irrelevant [AF17]. This selective change in sensitivity or perception of a new stimulus is known as "visual adaptation". The effects can occur on a time scale, being able to follow patterns of change to the stimuli, which reveals strategies of visual coding. For example, comply with the function of increasing the efficiency of storing and retrieving information until new information is highlighted [Web15].

Within the literature on NPR, previous research has been found in perception studies, which aim at the inference of perceptual rules through exposure to stimuli [Her10]. Three types of quantitative data are often considered in evaluating visual perception in measurements [LMD13]: study content scoring, response times for recognition and memorization, and eye-tracking (a technique that records eye movements, in particular, its central structure, the pupil). In this regard, the pupil is identified as an organ whose behavior can be interpreted as an indicator of emotional response to stimuli, being able to dilate before the presentation of affective images [SOBENG16].

With the help of vision data collection using eye-tracking, two studies on reaction to NPR images have been performed [MML11] [MML12]. The results showed an affective or emotional flattening effect, that is, a difficulty or inability to generate emotions, verifiable through the dimensions of activation of these and of pleasure. Compared to the original image (PR), the images filtered with NPR produced confusion, distraction and loss of interest in the users, making it difficult to interpret them. However, within the tested NPR algorithms, a photo-abstraction method originated affective responses closer to the original image, since it retained some essential details. In accordance with the above, it is pointed out that ambiguous images turn out to be more interesting, due to a natural desire to resolve uncertainty and seek clarity [Her10]. In contrast, vague images do not turn out to be very interesting. This is a good precedent for NPR when not overused.

2.3. Visual perception studies using eye-tracking

Within the production of 3DS images through NPR algorithms, all have been shown to be consistent, although some filters have produced more natural images than others [NAK13]. Differentiating more or less "natural" images implies, in practice, establishing that there is a range of possibilities within this combination. The depth perception "increased" by the 3DS could be combined with various visual qualities produced by NPR. This spectrum of visual styles, simultaneously NPR and stereoscopic, allows us to imagine different uses depending on the aesthetic needs of each audiovisual production.

In this regard, it is argued that the quality of 3DS images improves with the recognition of depth information [BLCC08]. Now, to estimate and calculate depth, the ability to understand the organization of the different elements that make up the scene and whether the depth information is realistic or plausible [LRBL12] tends to be considered. It should be noted that the quality can be evaluated with objective parameters, but there is a possible range of variations within what is considered a good 3D quality. This is especially evident when we explore the combination of NPR and 3DS.

There is a 3DS quality baseline, but there are different qualities of NPR textures that, not altering depth perception, add different aesthetic qualities to the same objects and backgrounds in the scene.

2.4. Ocular reactivity to NPR in stereoscopic cinema

In the evaluation of multimedia audiovisual quality [AF17] tends to consider human, contextual and technological influencing factors of the system, as well as the damages on the devices handled by users, such as blurring, that is, loss of information spatial or sharpness. However, it should be noted that the end user has been considered as the most prominent factor. This implies that measures of quality of experience (QoE), that is, user satisfaction, and quality of perception (QoP), the ability to analyze, synthesize, and assimilate content are preferred, instead of just the quality of service (QoS), the technical part of the medium. Within the research area on the problems and disadvantages of 3D, there are findings [UW11] [Sol13] that indicate that both images and 3D films, compared to 2D, are more effective in causing discomforts such as visually induced motion sickness [KDK10]. For this reason, the technology must be handled gently and applied carefully to ensure comfort and good performance for users [MHG14]. With poorly administered stereoscopy it could cause eyestrain for the viewer. Also, the symptoms of fatigue can manifest themselves consciously, such as headache and tiredness, or unconsciously, as disturbance of the oculomotor system [DB10]. There are side effects from stimulus perception that could happen in virtual reality [WS15].

2.5. Ocular reactivity to NPR in stereoscopic virtual reality

Virtual reality (VR) is a computer system for representing images and simulating real-life environments. Through display devices, such as the head-mounted display, either monocular or binocular (stereoscopic), it creates in the user the feeling of being immersed in the created world. Concepts such as: immersion and presence are derived from this VR technology. Remember that, immersion is that capacity of the system itself to involve you, and presence, the sensorial fidelity that it offers the user to make him feel in the place [SW97].

The VR user can experience, with the same level of immersion, different levels of presence, since it depends on a variety of human, technological and contextual factors [AF17]. In relation to this, the evidence [Toc16] shows that visual realism can positively affect presence (in VR). However, realism may not be determined by visual fidelity but by psychological fidelity: the extent to which the stimulus presentation evokes the type of physiological or emotional response one would experience in real life [W2]. Presence does not imply visual realism [SW97], or an experience with a representation style close to photorealism [UDWB16].

3.1. Introduction

In this chapter we present the methods used in our experimental research. 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 1920x780 pixels.

3.2. 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. 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. [MML12] but we chose a non 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 with the same duration.

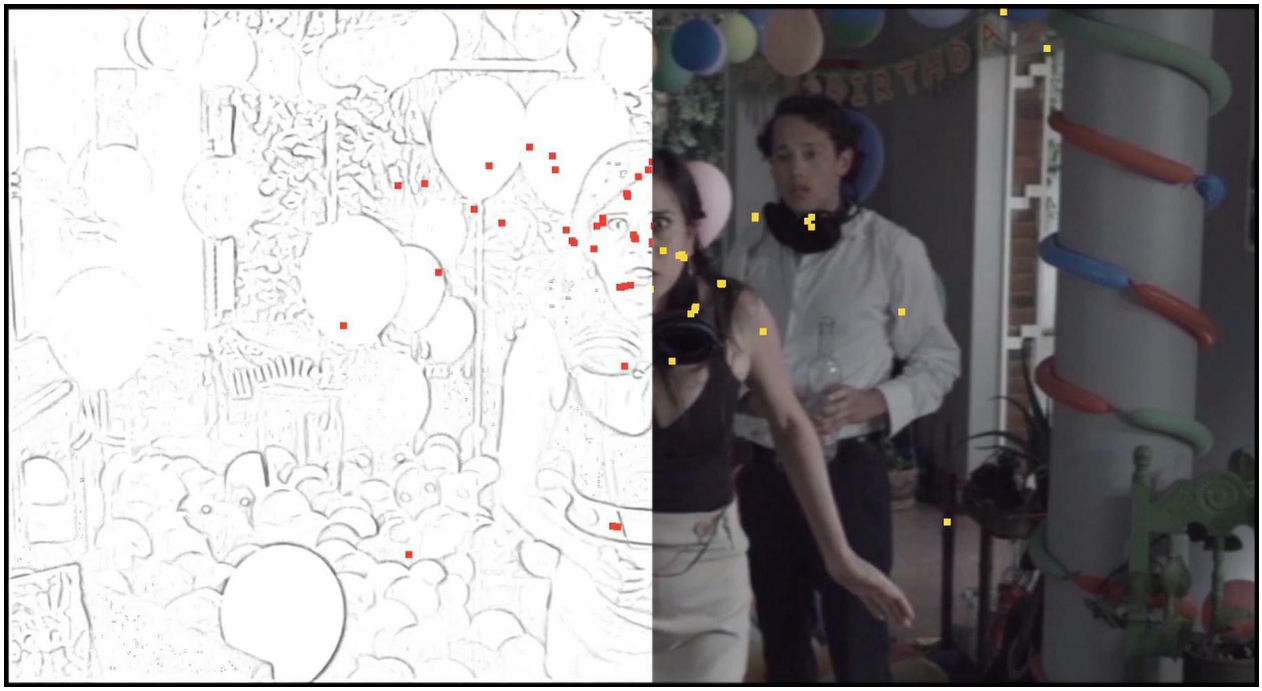


Figure 1. Styles used in the experiment: non-photorealistic (NPR) vs. photorealistic. The 27 dots, red and yellow, correspond to gaze fixation points of 27 spectators. The difference of color is just due to the mix of two visual styles in the same image.

3.2.1. 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 [SH06]. 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. Many NPR effects operate with randomly adding or removing lines or texture details such as hatching [USS11] or watercolor rendering [RKD10]. 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 [USS11] and also an improved stereo coherence [RKD10]. 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 [HI16] which would allow a limited use of object space methods. Also, the depth maps generated by these methods often fail to meet the quality requirements of cinematography.

There are many NPR algorithms designed to create the abstraction on the image [KD08] [WOG06] and thus allowing to enhance or guide perception [RD09] [SD04]. 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. [KCWI13] there is no mention about stereoscopy. Most stereoscopic NPR approaches include line drawings [KLKL13], watercolor rendering [RKD10] or artistic stylization [NAK13], however, we are interested in abstraction methods. As an edge enhancement method, we used the flow-based, extended difference-of-Gaussians (DoG) [WKO12] 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 [WOG06]. Texture details are removed by the iterative application of the bilateral filter, whereas color complexity is reduced using a smooth quantization step [WOG06] that avoids sharp transitions and also greatly increases temporal coherence, making the method applicable to videos or other real-time applications such as computer games [MSGS13]. Based on these references, for the purpose of our study, we decided to use a realtime 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 [Vid1] and [Vid2].

3.2.2. 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.

3.3. 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 the videos.

	Rendering	Sound
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 thesis 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 future work of this thesis 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.

3.4. 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.

3.5. Data processing

The data processing was carried out in collaboration with Larry González, a computer science engineer from the University of Chile and a PhD student in Artificial Intelligence at the Technische Universität Dresden, Germany.

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. Afterwards, 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 contained identifiers for user, video, frame, and the coordinates of the point of interest (x and y). It was composed of 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 [Hot51] present in the R package 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 [Lin91], defined by the following equation, as a measure of dispersion, where $P(x_1)$ is the probability of occurrence of event x_1 .

$$H(X) = -\sum_{i=1}^{n} P(x_i)\, log(P(x_i))$$

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

3.6. 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 ocular fixation. 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 in terms of ocular fixation data, we identified two sets of results (see Figure 6 in page 53). In a first group composed of 8 episodes, where no significant differences were observed between the photorealistic (PR) and non-photorealistic (NPR) videos. In the second group, also composed of 8 episodes, we observed significant differences between PR and NPR videos. This data allows us to statistically characterize the ocular behaviors but we needed a qualitative approach to describe the audiovisual content, in terms that allow us to compare the 16 identified episodes. We make this qualitative description using two types of descriptors: “cinematic procedures” and “cinematic events”. On the one hand, we define the “cinematic procedures” as those factors potentially present in any cinematographic content: editing cuts and camera movements. On the other hand, we define the “cinematic events” as elements placed inside the filmed scene: fluctuations of characters and objects within the scenic space.

3.7. Conclusive remarks

In summary, combining quantitative and qualitative criteria, we analyzed the data and established that there are possible correlations between the cinematic content and the observed ocular behavior.

4.1. Introduction

Our interdisciplinary research is dedicated to exploring the boundaries of stereoscopic filmmaking from an unusual viewpoint: we aim at creating 3D non-photorealistic cinema which allows conciliating a stereoscopic pre-visualization that is oriented to ensure visual comfort with concept tests of NPR applied to a 3D film. In this chapter we describe the role of pre-visualization in stereoscopic cinema and our preliminary observations and experience of combining 3D cinema with non-photorealistic rendering approaches, from the filmmakers' point of view.

The history of cinema is indissolubly linked to technological development and photorealism. Since the beginning of the digital era, technological advances have not ceased to push further the boundaries of photorealism. A good example is the huge development of special effects and 3D animation, both oriented to reach the empathic effect of a photographic image. So, in general terms, photorealism has aesthetically dominated the history of cinema through the different stages of its technological development.

In the case of 3D cinema, photorealism is even more important, at least for two factors: the promise of sensorial hyperrealism is the main commercial argument of 3D experience in movie theaters, and at the same time, it is a sort of universally accepted standard for producing a visual comfort in the viewer. Thus, hyperrealism and visual comfort seem to be two inseparable characteristics of contemporary 3D cinema, both resting on the idea of photorealism as a conventional premise for stereoscopic filmmaking. We will question this premise by exploring the possibility of a non-photorealistic 3D cinema, in a two-step research.

The first step will consist of evaluating the use of non-photorealistic rendering (NPR) filters in 2D cinema on two types of users: general public (university students) and specialized public (film post-production professionals). The second step will consist of exploring the perceptual and technical problems of using NPR in stereoscopic cinema: on the one hand, to understand the previsualization method as a way to guarantee visual comfort for the whole stereoscopic content, and on the other hand, to identify the technical and perceptual challenges of combining stereoscopic visualization with NPR.

Even though believable 3D relies on a realistic rendering model, realism is only one of the many artistic styles of storytelling and expressing emotions: choosing the proper visual style is an important aspect of art [McC94]. So, if we think of the aesthetics of 3D cinema as a spectrum of possibilities ranging from photorealism to non-photorealism, this idea expands the possibilities of cinematographic creativity. In this chapter, we show experimental results of NPR applied to 2D cinema, and preliminary observations of NPR applied to 3D cinema. Thus, we aim at exploring the limits of merging the expressive possibilities of NPR with the narrative use of stereoscopy.

In Section 4.2, we describe visual comfort as the main perceptual challenge in 3D cinema and previsualization as a tool to achieve it. In section 4.3, we describe the framework we have developed and used to test the role and applicability of NPR to stereoscopic cinema. Section 4.4 presents results of NPR for 2D cinema. In Section 4.5, we explore a two-pattern method to compare samples of stereoscopic contents processed with NPR filters. Finally, we conclude our chapter in Section 4.6.

4.2. Visual comfort as a perceptual challenge for 3D cinema

Constant "depth scanning" is a natural characteristic of human vision. But even if the principle of binocularity is shared in stereoscopy, the 3D cinema of the 21st century still produces ocular discomfort, dizziness or headaches in many people [UH08]. About this aspect, in 2014 a French public agency published a series of recommendations on possible damage detected in children under 6 when exposed to stereoscopic content [Ans14], in order to limit risks on children's health. The French agency[2] that produced this report states, however, that "given the lack of information on exposure to these technologies, the Agency considers necessary to promote new research".

The "know-how" for creating high quality stereo motion pictures with stylized graphics is very limited. Even though it was shown that the mixture of stereoscopy with NPR can break 3D space perception [GW02], there has been little research on how to create visually pleasing NPR images or videos. Recent research focused on specific sub-problems such as line drawings [KLKL13], painterly rendering [SG04] [NAK13] or simulating film grain effects [TDMS14]; however, a general framework to combine arbitrary styles (including post-production work that is applied in image space in 2D) with stereoscopic 3D visualization is yet

[2] The French Agency for Food, Environmental and Occupational Health & Safety (ANSES): https://www.anses.fr/en

to be established.

In this chapter we merely scratch the surface of this topic: we experiment with screen-space NPR stylization methods that are applied independently on the two images corresponding to the two eyes. Among other issues, we are interested in how stylization influences the filmmaking procedure, including the quick generation of previews or pre-visualization ("previz") with a set of different styles. It is not enough to just create solutions that optimize 3D conventional processes (such as geometric or light correction applied to stereo pairs); it is fundamental to experiment with the language of 3D cinema and, in the same process, analyze the user's responses to formal innovations. To achieve this, we have shot a stereoscopic film following the rules of a "secure" stereography, which means using pre-visualization as a tool to prevent disturbing transitions between scenes with different depths. This film would be used to test with NPR processing.

4.3. The role of cinematic factors in visual comfort

The experience of filming in 3D is very didactic to understand the concept of visual comfort from an empirical perspective. In principle, the inconvenience is ocular [SH09] [LFHI09]: eye fatigue results from the conflict between fixed accommodation on the screen (where we focus) and mobile convergence in scenic 3D space (where we look). In the 3D experience, although we can get used to this process after a few minutes, the quality of the film does the rest. The 3D quality of a film responds to decisions of "mise en scène" because they simultaneously involve several components of cinematography (photography, editing, post production). We shall call these the "cinematic factors" of eye fatigue. If we compare the "physiological" and "cinematic" factors of eye fatigue, we should assume that, as medical research, optics and neuroscience do not submit new evidence on how to reduce the impact of physiological factors, our efforts should focus on cinematic factors. By properly using conventional resources of 3D cinema, we could produce films that are comfortable for the average viewer. One of these resources to avoid eye fatigue is the stereoscopic preview or 3D storyboard. The experience of shooting in 3D, which we describe below, incorporated the use of stereoscopic preview as one of the cinematic strategies to define a shooting plan aimed at visual comfort.

In July 2014, we shot a 3D film at the residence of Spanish filmmaker Luis Buñuel in Mexico City, based on a three-dimensional model of the house. This model allowed identifying the camera axes with greater visual depth, so that the shooting could take full advantage of depth

perception in that house.

Once the script and the model of the house were completed, Frameforge [Fra] was used to simulate the material aspects (interior and exterior of the house, furniture, lighting, characters), optical factors (photographic focus, stereoscopic variables) and cinematic factors (fragmentation of history into scenes and shots). This stereoscopic preview was useful to generate the optimal amount of material for edition, which is very useful to solve problems in 3D post-production. Then, there is the aesthetic role of post-production: what kind of "look" is better adapted to the fantastic spirit of the story. That was the beginning of our NPR real-time software testing. We wanted to develop a tool that was useful for matching the story and its aesthetics: to merge the expressive (NPR) and the narrative (3D) dimensions of the film. We knew that the natural environment of 3D cinema is photorealism, but the surreal context of the story gave us the aesthetic justification for a non-photorealistic treatment. The next step was testing the NPR filters in a 3D film conceived with cinematographic principles based on visual comfort [LHCLLC11].

4.4. Test framework of NPR effects for cinema

From a technical point of view, NPR methods can be classified into object space and image space (or screen space) methods. Although in our case study the 3D model of the scene was available, in general, only film shots were available and so we chose image space methods for stylization. These approaches work with 2D image streams and therefore can use only color and texture data. However, we may assume that limited geometric information is available in the form of depth images, which may be extracted from the stereo images or directly captured using an additional depth sensor during film shooting. Thus, in addition to standard image processing methods we also considered depth-based effects that are also calculated in image space, but pixel data may correspond to depth.

Among image space NPR methods, we looked for ones that allow interactive performance. This may seem contradictory, as rendering methods used by cinematography are traditionally performed offline. On the other hand, we intended to include this stylization framework into the fast previz stage of film production where many different styles are tested rapidly, possibly during shooting, and thus performance is favored over high quality. Based on the preliminary results shown by the previz stage, high quality offline (possibly manual or semiautomatic) methods may be developed or selected in later production stages.

Our previz tool is implemented as post-processing effects, using the Unity game engine [Uni]. We implemented a stand-alone video editor software in the same platform that supports various parameterized NPR effects. These effects execute basic image processing algorithms on the GPU and are capable of realtime performance. Thus, users can see the original shot and the immediately computed stylized results both at once. The selected effects are based on two principles: artists enhance relevant details and at the same time simplify the image by mitigating irrelevant details. Relevant details are emphasized by drawing lines (i.e. applying edge detectors).

We used the flow-based, extended difference-of-Gaussians (DoG) filters proposed by Winnemöller, as these were shown to produce aesthetically pleasing results [WKO12] and can simulate various effects such as black and white or colored pastel. It is also related to edge detection by approximating the Laplacian of Gaussian (LoG) filter, which is equivalent to blurring the input with a Gaussian filter and then applying a Laplacian, i.e. second order edge detection. The result of the DoG filter is thresholded: smooth thresholding is used by applying the tanh function in order to produce aesthetic results [WKO12]. To avoid noisy responses, smoothing along the flow field is used. First, we compute the smoothed structure tensor (i.e. the standard structure tensor for color images, blurred with a Gaussian filter), from which the gradient and tangent directions are extracted, similarly to [KD08]. Then, in a second pass, line integral convolution that follows the edge tangent flow is applied. Line parameters such as color, width and smoothness are user-controlled parameters in our system; the parameters correspond to the reparameterization of the DoG filtered as proposed by Winnemöller in [WKO12]. Black and white contour enhancement using the flow-based extended DoG filter is referred to later on as “Added contours”. We also defined an effect that takes the original pixel color and uses it as edge color, referred to as “Colored line drawing”. Additionally, with proper threshold parameters, the DoG filter can produce pastel-like effects [WKO12], which we consequently named “Pastel” and “Colored Pastel” (see Figure 2 in page 40).

Image simplification methods consist of lowering image complexity in terms of texture details and color details. To reduce texture complexity, we used an extension of the flow-based implementation [KD08] of the bilateral filter [TM98] [WOG06], which is an edge-preserving smoothing filter. In our case, the bilateral filter is a product of two Gaussian filters; one is applied in the spatial domain, and the other is computed in the intensity domain. The intensity-dependent filter component ensures that neighboring pixels that are placed on the same side of a step-like signal as the center pixel have greater weight, while pixels from the

other side of the edge contribute less to the filter output, better preserving the edge. The amount of blur is controlled by the variance parameter of the spatial domain Gaussian filter, while the amount of details kept is determined by the intensity domain variance parameter. In order to avoid color-bleeding artifacts, the bilateral filter is applied in the CIE-Lab color space [TM98]. Originally, the bilateral filter is non separable, and thus expensive to compute. In real-time applications, usually the flow-based approximation [KD08] is used instead, which applies a one-dimensional bilateral filter along the gradient flow, and then another one-dimensional bilateral filtering on the tangent flow. Similarly to the flow-based DoG filtering, the gradient and the tangent flow directions are extracted from the smoothed structure tensor. The bilateral filter is usually performed iteratively several (2-5) times to produce visually appealing results. Color complexity is lowered using luminance quantization, similarly to [WOG06]. We refer to the output of the application of the bilateral filter and luminance quantization as a "Simplified" image. Another way of lowering color complexity is to reduce image saturation ("desaturation") in HSV on HSL color spaces. The combination of these effects together can simulate different artistic styles. Additionally, the level of abstraction is parameterized in each of the effects: i.e. the line thickness and density in edge detection, the strength of details that are kept in texture simplification and the amount of desaturation. This allows us to guide the viewer's gaze [RD09] [MSGS13], as well as to create the illusion of depth. The parameters corresponding to the level of abstraction may depend, in the first place, on camera depth, as a way to show objects in full detail in the foreground and as an abstraction in the background. Another typical use is to define the level of abstraction based on the radial distance from a particular point on the image, which guides the viewer's gaze to this particular point.

Effects are applied on the two stereo images independently —without modifying the stereoscopic parameters. This is known to break 3D perception [KLKL13]. However, in our case this is less noticeable for several reasons. First, contours are generated on each-eye basis, which was shown to be capable of avoiding binocular rivalry [KLKL13]. In most styles (except for line drawings), there is interior information between lines, mitigating binocular rivalry caused by line segments that are seen only by one eye. Additionally, lines are not textured, and thus line style is coherent. Finally, image simplification methods usually aim at removing less relevant details while preserving relevant ones. Thus, these approaches remove details instead of introducing new ones.

This means that most of those details are present in the output image, which were also visible in the original, photorealistic shot; everything else is smoothed out.

Our preliminary user studies showed that the consistency of the two images can be high and thus disturbing artifacts may often not be present: 3D illusion is not affected by some of the NPR effects. We note that the effects are intended to be used for previz, and thus an offline stylization method that is used to render the film in its final form should be very carefully designed in order to achieve perfect 3D sensation.

4.5. NPR rendering for 2D cinema: experimental results

An experimental phase of the research was carried out in order to identify and analyze how cinematic perception in 2D is affected by NPR. Given the cultural centrality of our habit of 2D cinema, we considered a priority to compare three "extreme" or "polar" (clearly distinguishable) types of NPR filters. We worked with a young audience of university students, characterized by high audiovisual consumption. The experiment was conducted in October 2014 in Santiago, Chile, and it was conceived to identify the conditions in which visual abstraction of NPR processing can affect the narrative understanding of a 2D film.

Figure 2. Three samples of NPR filters: Native frame (left), COLOR COMIC (center-left),PASTEL (center-right), PAINTING (right).

We applied three NPR filters, COLOR COMIC, PASTEL, PAINTING (Figure 2), in a 7-minute fragment of the feature film "Las Niñas Quispe" (Dir. Sebastián Sepúlveda, winner of best cinematography at Mostra di Venezia 2014). Then we organized a screening in a movie hall of these three NPR versions, divided into three groups of 10 students each, from a total of 30 college undergraduates of cinema and journalism, ages 19 to 22. After each screening, we combined a quantitative methodology (survey of 25 questions) and a qualitative questionnaire (three focus groups of 10 people each). The first part of the survey allowed us to validate the baseline of a general high audiovisual consumption on that group of college students. The

second part consisted of questions divided by themes: space-time perception, emotion identification, recognition of characters and backgrounds. The focus group results allowed us to build an analytical reading of the survey's results.

For the three groups of students, the qualitative consensus was that character recognition, more than backgrounds or objects, is the most important factor to evaluate the impact of NPR on film narrative. This result could suggest that the NPR do not affect depth perception because character recognition remains significant even in the "extreme" filters (Pastel and Painting), where the backgrounds and foregrounds seem to be combined in similar textures.

As we can see in the quantitative results (Figure 3, left), the morphological perception of characters (facial features, sizes, textures) was clear in NPR1 (Color Comic); the vertical axis indicates the number of mentions. In NPR2 (Pastel) greater importance was given to voice as a differentiating criterion, due to the general darkness. However, in NPR3 (Painting), voice becomes the main element of recognition, leaving the morphology in the background. We concluded that one reason is the visual disturbance of the predominance of white, which seems to decrease visual perception and encourage auditory perception. Colors (blue component) was not mentioned as a recognition criterion, and a combination of factors, labeled as Others (violet component) was only mentioned in NPR1 and NPR3.

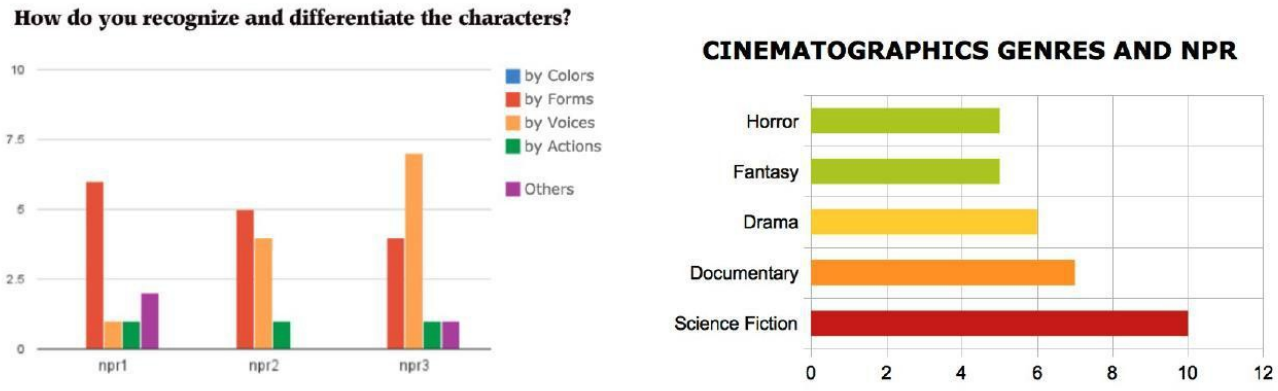


Figure 3. Left : comparative factors for character recognition in NPR1, NPR2 and NPR3. Right: the point of view of film post-production professionals in Chile, about the potential use of NPR in cinematographic genres.

These results suggest that we do not need so much detail (as we found in photorealism) to understand a film narrative, because, in fact, much of the information comes from character identity, produced by the combination of body movement and voice. Therefore, when using NPR, subjects are able to abstract, relate and follow a story, even when we see more blots and stains than sharp facial expressions. NPR seems to bring a new atmosphere to the story,

without affecting its dramatic understanding. The use of “extreme” filters, two of them clearly distant from the native realistic picture of the film, stimulates new interpretations of the characters emotions and intentions and the narrative importance of their environments. In short, understanding the story does not seem to be affected by NPR filters. The main conclusion to be drawn from this experiment of non-photorealistic 2D film lies in a dual dynamic generated by NPR: general narrative comprehension is unaffected but, by modulating the morphological aspects of characters and by increasing the perceptual importance of sound, new approaches to the film intentions are generated. The same story seems to produce new interpretative variants in its viewers. This seems to happen in 2D.

The second result, based on a panel of 12 professionals of film post-production, is related to the use of NPR filters in cinematographic genres (Figure 3, right), where the X axis indicates the number of professionals that mention a genre There was a genre in which professionals were more open to seeing NPR: Science Fiction. This suggests that post-producers imagine a possible use of NPR in movies with non-realistic stories. But the surprise comes with the second genre (documentary), generally more associated with realistic characters or landscapes.

If we confront the results of both groups (students and professionals), we find no contradiction: if NPR filters do not affect character recognition, the stories will be correctly understood. However, to explore passing NPR from 2D to 3D cinema we must remember the main perceptual difference between monocular and binocular vision: spatial understanding. In 2D, a singular visual scanning is enough to understand instantly the spatial nature of the scene, but in 3D, visual scanning is essentially comparative: an ocular backward and forward movement that could be called “depth scanning”. Therefore, even if the characters remain the most important factor of NPR perception in 2D, it will be necessary to evaluate how this constant “depth scanning” of 3D perception could modify the impact of NPR.

4.6. Comparing samples of 3D frames processed with NPR

As we said before, “depth scanning” is an attribute of 3D perception. Buñuel’s house, used for the shooting, was filled with objects in order to increase depth perception in the filmic space. We were also sure about the visual comfort of the photorealistic 3D content, that means, before the NPR processing. We also had the NPR software, so we needed to prepare the concept test of processing a scene with different NPR filters. And the final step —which we shall present as future work— should be the user-test of NPR 3D with viewers. So, we created a two-pattern method to compare the NPR 3D samples. Instead of defining depth perception by optical principles, we first developed a comparative method of NPR effects with a specific validation criteria: the expressive and narrative needs of the film. We chose a scene with six different layers of depth, where the narrative intention was the sadness felt by the female character.

The comparison of the resulting images after NPR processing, analyzed with anaglyph glasses, was described in terms of 3D quality (perceived layers), and related to the impact of NPR on stereoscopic illusion (whether it keeps the binocular depth or not). We know that anaglyph is not the best way of visualizing 3D (polarized glasses are the cinema standard), but at the same time, all 3D contents available on the Internet are in anaglyph format. Currently, both kinds of glasses coexist, but in this exploratory phase, anaglyph visualization was sufficient. We must assume a general lack of “3D culture”. Almost everyone has seen 2D films. But in adult ages, even among cinema students, the experience of watching 3D films is still rare. We should not forget these differences in cinema backgrounds for our future NPR 3D user tests. That is why we started with two intuitive criteria for the comparison: (a) 3D quality, and (b) depth perception. We had to conceive intuitive notions of 3D that could be easily expressed in a social context of users with no “3D culture”.

	Native Shot	Non-Photorealistic Shots						
	(A) Photo-realistic	(B) Colored pastel	(C) De-saturated color	(D) Saturated color	(E) Added contours	(F) Simplified	(G) Pastel	(H) Colored line drawing
3D quality	clear layer separation	layers with ghosting	layers with ghosting	clear layer separation	clear layer separation	clear layer separation	no layer separation	no layer separation
Depth perception	binocular depth	monocular depth	weak binocular depth	intense binocular depth	realistic binocular depth	unrealistic binocular depth	monocular depth	monocular depth

Table 2. Comparison of different NPR filters applied to a stereoscopic frame.

We see in Table 2 the "native" photo-realistic shot compared with seven non-photorealistic shots as a preliminary basis for a user test that will be presented as future work. In columns (B) and (G), pastel refers to the pastel-like effect from [WKO12], in (C) and (D) columns (de)saturation means (lowering) increasing the saturation of colors based on the radial distance from the center of the image, in (E) column the results of edge detection were added, (F) column is an image abstraction method based on [KD08], while in (H) column the result of an edge detector was taken and the original input was used as edge color.

From the point of view of perception, the choice of these seven non-photorealistic filters was oriented to produce a continuous scale of proximal types of depth, contrary to the 2D NPR experiment, in which we chose two "polar" filters and just one proximal filter to the "native" frame. Once the scene was processed with NPR filters, we could work on a questionnaire based on two patterns, that could be summarized in two kinds of questions: a) about "3D quality": do you see "edge ghosts" when you compare the depth levels of the image? b) about "depth perception": which samples seem to appear to you as a normal 2D image?

A simple comparison of perceptions among the research team (we all have a different binocular vision) revealed a comparative panorama of which NPR filters produce different kinds of 3D illusion. For instance, NPR samples that reduce the stereoscopic illusion by producing a monocular depth perception (B, G and H in Table 2) are less suitable for stereoscopy. Something similar may be concluded for those samples that create a defective stereoscopic illusion (B and C in this case), by producing "ghosting" around the edges. Combining these two intuitive criteria we could select the most suitable NPR effects in terms of visual comfort (D, E, in F in Figure 4).

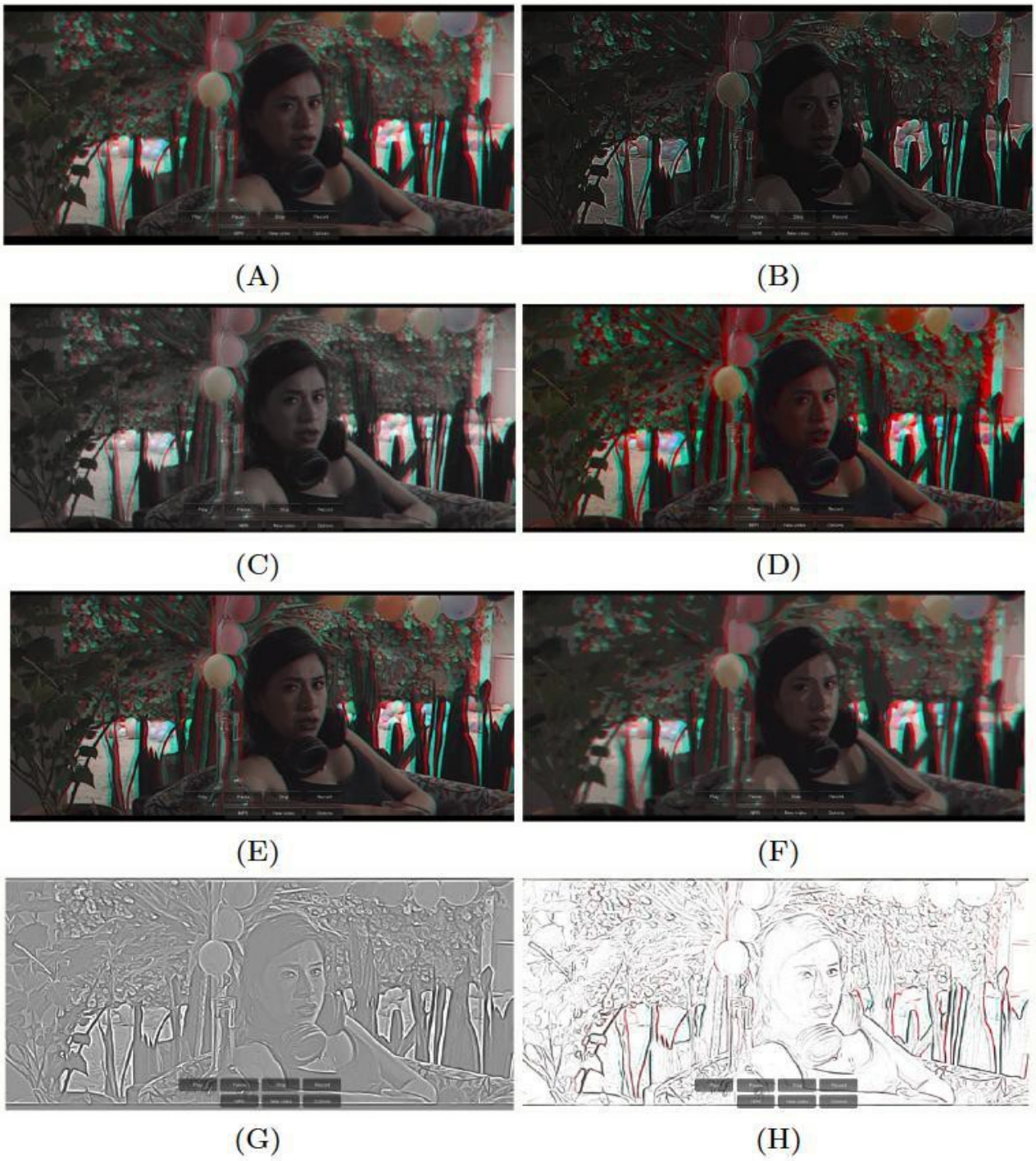


Figure 4. Photorealistic shot (A) and non-photorealistic stylized shots (B–H). NPR samples that reduce the stereoscopic illusion by producing a monocular depth perception are B, G and H. The samples that create a defective stereoscopic illusion, by producing "ghosting" around the edges, are B and C. Combining these two intuitive criteria we could select the most suitable NPR effects in terms of visual comfort are D, E, and F.

Even if these preliminary results are just a first intuitive step in our methodological path towards a systematic experimental strategy, we can already identify some ideas to consider for our future work: a) we could use optical solutions (as eye-tracking devices) to corroborate the verbal identification of "ghosts" in the figure's edges; b) for a qualitative approach, we should consider intuitive definitions of "3D quality" and "depth perception" if we want to consider the lack of "3D culture"; and c) also for a qualitative approach, we must define a set of descriptive attributes that would be useful to compare and establish differences between the expressive contribution of each NPR filter in the same 3D scene.

4.7. Conclusive remarks: exploring the limits of photorealism in cinema

The use of NPR in cinema means a challenge to the global tradition of photorealism. The common sense of our visual culture tells us that NPR and stereoscopy are destined to be separated. This thesis suggests that this combination is not a perceptual contradiction. Our experimental results in 2D cinema indicate that narrative comprehension is not significantly affected by NPR. In 3D cinema we conducted a trial study that produced preliminary observations. Our current work indicates that a combination of the expressive quality of NPR with the immersive effect of stereoscopic cinema could produce a new form of augmented narrative.

5. Tracking gaze behavior: attentional reactivity to NPR in 3D cinema

5.1. Introduction

Our research is about the effects of non-photorealistic rendering in the perception of stereoscopic cinema. We conducted a study with 27 participants, using eye tracker to evaluate ocular behavior during the free viewing of a stereoscopic film. Using eye tracking technology to study ocular behavior during cinematic immersion, we used two types of ocular data: ocular fixation, to study the evolution of visual attention, and pupillary reactivity, to evaluate the emotional response. We correlated attentional and emotional behavior with the evolution of visual entropy. Our results leads us to believe that, facing a highly abstract non-photorealistic scene, the gaze seems to react under the effect of a perceptive anxiety. This ocular response seems to spontaneously assign greater sensorimotor resources to the analysis of the character's behavior, transforming his face into the main key to analyse the information and sense of the global scene. This behavior appears to us as an adaptive response to a visual environment endowed with abstraction and cognitive uncertainty. Our results could be useful to conceive new rendering tools for the post-production of stereoscopic films and interactive virtual reality experiences, in fields like academic research, arts & culture, education & museums, medicine & rehabilitation, training simulators.

How do we perceive the visual abstraction of non-photorealistic rendering processing in a 3D movie? Using an eye tracker to produce and analyse gaze data, we seek to evaluate if there are possible correlations between non-photorealistic post-production and a greater immersion in the narrative of a 3D movie. In this context, we use the expression "augmented film narrative" to evoke the hypothesis of a differentiated gaze behavior, triggered by the use of NPR, that could produce a higher immersion on the film narrative. Our results indicate that NPR processing, compared to photorealistic style, is globally correlated with lower levels of visual entropy. The visual abstraction of NPR stimulates a higher gaze concentration in less elements of the scenes, in a temporally dynamic behavior which appears to be opposite to the temporal monotony of ocular response to photorealistic scenes. In the context of our case study of visual abstraction and film perception, we used this entropy as an empirical measure for a better understanding of attentional synchrony during the free viewing of a non-photorealistic content. We also discovered that photorealistic and non-photorealistic styles produce

common ocular responses to cinematographic procedures such as editing and camera movements, but NPR processing triggers a significant increase of fixations in the bodies and faces of the characters. The face works like a magnet for the gaze with greater intensity in the videos with non-photorealistic processing.

We combined neurobiological and sociological explanations for this phenomena. Facing a non-photorealistic scene, characterized by a correlation between visual abstraction and cognitive uncertainty, the gaze seems to react under the effect of a perceptive anxiety: an ocular behavior that can be described as a limited exploration of the scene and a greater focus on the most expressive elements of the scene, such as bodies and faces of the characters. This behavior appears to us as an adaptive response to a visual environment endowed with abstraction and uncertainty. This perceptive anxiety seems to spontaneously assign greater sensorimotor resources to the analysis of the character's behavior, transforming his face into the main key to analyze the information and the emotional tone of the global scene. In this sense, the corporal elements seem to be used as perceptive keys to interpret, inferentially, the global meaning of a non-photorealistic scene. In an abstract visual context, this perceptive anxiety could trigger an ocular behavior defined as an adaptive strategy to reduce cognitive uncertainty as quickly as possible. A better understanding of how this effect of cognitive uncertainty triggered by visual abstraction can be used to develop better NPR tools. We believe that these results could contribute to conceive the non-photorealistic post-production as a tool that allows enhancing the cognitive immersion of the viewer in certain types of stories and film genres (fantasy, suspense, science fiction), using abstraction and strangeness as an aesthetic effect to enhance the film narrative.

5.2. Related research

Gooch et al. develop a study to evaluate the effects of NPR on perception in the simulation of immersive spaces [GW02]. This study proposes a relevant conclusion regarding theoretical studies about NPR: whether in expressive or functional mode, NPR seems to generate an effect of saliency in chosen elements over the other elements of an image. Halper et al. [HMHLS03] conclude that photorealistic rendering (PR) can effectively render objects and images in non-realistic styles without influencing feature binding processes necessary for basic object identification. The NPR processing then seems to favor the salience of certain visual elements without hindering the analysis of their significance within a scene. Using the eye tracker as a measuring tool, Cole et al. [CDFKMS06] show that NPR rendering can be

controlled to affect gaze position. Another relevant precedent appears in a study carried out by Santella et al. [SD04], which is used to measure effects produced by processing images with NPR filters, about how they draw attention. They demonstrate that it is possible to utilize this tool to obtain relevant information about the interaction of the gaze with different levels of abstraction in an image. These studies show that visual abstraction has attentional effects, but we do not find studies focused specifically on evaluating these effects in the 3D cinematographic field. More specifically, NPR seems to focus attention but there is not enough evidence concerning this perceptual influence in the case of stereoscopic cinema. Our interest consists in evaluating the relevance of this knowledge on the attentional and perceptual impact of NPR using the eye tracker as a tool to describe the ocular behavior. The relation between gaze behavior and narrative film comprehension has been analyzed by Loschky et al., identifying film editing as a crucial cinematic procedure for attentional continuity [29]. Factors that favor this attentional continuity have been studied regarding the perception of film editing, which describes the way in which perceptive phenomena occur. We are particularly interested in the concept of attentional synchrony, proposed by Smith [Smi12], because it provides the description of how the gaze of multiple viewers clusters around points of high motion. The study of 3D film perception concerns also the influence of sound design, because it has been shown that our perception of phenomena is often done through a combination of modalities —an effect known as cross-modality [SM93] [NGJT13] [VdG00]. Previous work has been made to evaluate the perceptive impact of sound in 2D and 3D films, proposing a multimodal correlation in film perception, where a greater degree of visual abstraction seems to be associated with a greater influence of sound to guide the recognition of characters [FMCOS17]. Nevertheless, there is insufficient evidence about the combined effect of NPR stylization and sound post-production in 3D cinema. Most of the descriptions of how visual perception of cinema works were made on the basis of photorealistic films. So, it is not clear if the described ocular behaviors are equally valid for non-photorealistic cinema contents. Thus, the goal of our study is to combine existing knowledge about NPR and film perception to contribute to the study of NPR in 3D cinema.

However, the dimensions of the question about the cognition of the moving image are more complex. On the one hand, the current knowledge of neuroscience, in the most recent enactive perspective, indicates that cognitive structures emerge from recurrent sensorimotor patterns that allow action to be guided by perception [VR04]. These structures involve different sensory-motor capabilities, which are part of a larger biological and cultural context [VR04].

The basic notion is that cognitive abilities are inextricably linked with a lived history [Var95]. Perception is constructed by the cognitive processes put into action where the subject uses his sociocultural memory as an interpretive key.

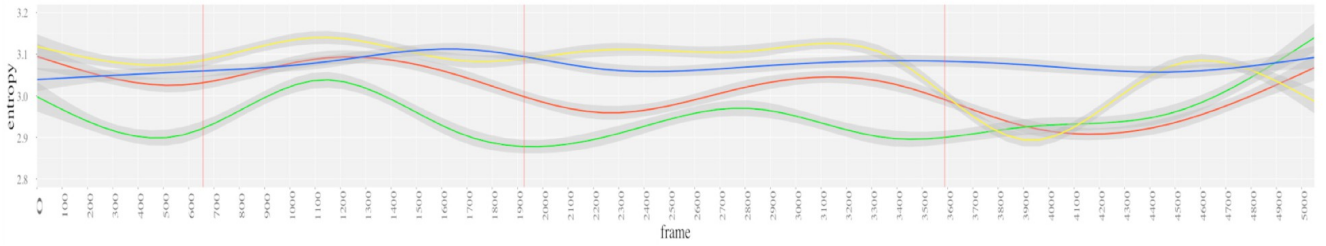
On the other hand, empirical research conducted in Chile regarding Internet browsing [Del12] demonstrates cognitive differentiations, in two intentional samples of 544 cases each, referring to young people and adults, where there are no differences by gender or social stratum, but by age. Based on surveys, a Cognitive Complexity Index was constructed in relation to how much information subjects can perceive and/or process at the same time. This index observes increasing levels of low complexity processing, as the age of the subjects increases [Del15]. At the same time attentional differences were detected: the younger focus more on the images, while the older focus more on the words. This evidence was validated with qualitative studies of navigation experiences, consistent with quantitative data from surveys. Then, from a theoretical point of view, both neuroscience and empirical research show us that there is not a single pattern of perception but that it varies according to sociocultural factors. In this multidisciplinary context, we can say that we have a plurality of cognitive processes, a plurality of forms of imaging (photorealistic/abstract) and sounds (narrative/immersive), and evidently, also, a plurality of tools to describe the processes of perception strategies. Given this diversity, we chose to describe a polar performance, that is, patterns of cognitive functioning that simultaneously imply the audiovisual format and the user's ocular behavior, but evaluating the relevance of the type of descriptors used. From this perspective, the objective of the research is to study a perceptive functioning pole associated with abstraction: the possibility of describing significant correlations between a non-photorealistic style and the cognitive processing of the user. Concrete utility is to determine if the post-production style can favor the user to focus on certain elements of a film to increase their degree of cognitive involvement in the film experience. It must be taken into account that users are not universal: they are embedded in a certain type of audiovisual culture. In our case, the study is delimited in Western culture, and is not necessarily valid for other cultures where audiovisual content can be structured according to other reading patterns, for example, from right to left, or from top to bottom, as in Eastern cultures. To do this, we developed two opposing post-production criteria: an image processing where photorealism and abstraction are opposed, and a sound processing where narrative audio opposes an immersive audio. In this sense, our research aims to make a small contribution to understand a larger problem, which we hope may have a heuristic value for future research.

5.3. Methods

Combining quantitative and qualitative techniques, we analyze the data to establish if there are possible correlations between the audiovisual content and the observed ocular reactivity. The detailed methods are presented in the Chapter 3 “Methods”.

5.4. Results

We established comparisons between the videos regarding their entropy (see Figure 5), 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 entropy: NPR-IS [green curve], NPR-NS [red curve], PR-NS [blue curve] and PR-IS [yellow curve]. 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.



Video	Visual Postproduction	Sound Postproduction	Curve color	mean(entropy)	Standard deviation(entropy)
4	photorealistic	immersive sound	yellow	3.0794	0.1789
3	photorealistic	narrative sound	blue	3.0787	**0.1861**
1	non-photorealistic	narrative sound	red	3.0192	0.2008
2	non-photorealistic	immersive sound	green	2.9611	0.2220

Figure 5. Comparison of NPR and PR videos in terms of visual entropy. The red and green lines, correlated

with non-photorealistic videos (1,2), show less entropy (visual dispersion), than photorealistic videos (3,4), in blue and yellow lines. It can also be noticed that the "native" video (3) is the less reactive in terms of entropy.

To compare the temporal dynamics of the entropy in the four videos, we complemented the qualitative analysis of the curves with the standard deviation as a measure of variability. We note a distinction between photorealistic videos (3 and 4), which show a lower attention oscillation than non-photorealistic videos (1 and 2). This is made even clearer in video 4, which differs from the other three videos because it seems to react less to editing cuts and the internal dynamics of the scene.

In attentional terms, video 3 presents a more monotonous behavior (less variations in time), which may be an indicator of a lower cognitive variability. Because we do not perform an analysis of pupil behavior as an indicator of emotional response, nor do interviews with users, we can not establish whether this lower cognitive variability can correlate or not with variations in the emotional involvement of the user in the film narrative.

5.4.1. Ocular response 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 permitted us to establish sound as a relevant factor in ocular response. Overall sound does not appear to be a trigger for a perceptual difference. Considering previous knowledge about the multimodal dimension of perception [VdG00], we consider it probable that this result can be explained by a wrong strategy regarding audio post-production in our experimental design. It will then be necessary to rethink sound post-production to evaluate this multimodal dimension in future research.

5.4.2. Ocular response to NPR post-production

General comparison. Three pairs of videos were identified with the presence of NPR filters. We compared them to establish if NPR generates significant differences in the same moments of the film. Each comparison generated a graphic where the X-axis represents the frames of the video and the Y-axis represents the p-value. The black curve represents the raw data of the p-value frame by frame, the straight horizontal line represents the threshold of 0.05 p-value. In blue, a softened curve is presented, which provides a local estimation with a window of 20 seconds. With the purpose of identifying the start and end points (In, Out) of each significant segment, the smooth curve went from being significant to non-significant and vice versa according to the established threshold (0.05). This resulted in the identification of 8 significant events in the distribution V1–V3, 11 significant events in the distribution V1–V2, and 6 significant events in the distribution V2–V4 (see Figure 6).

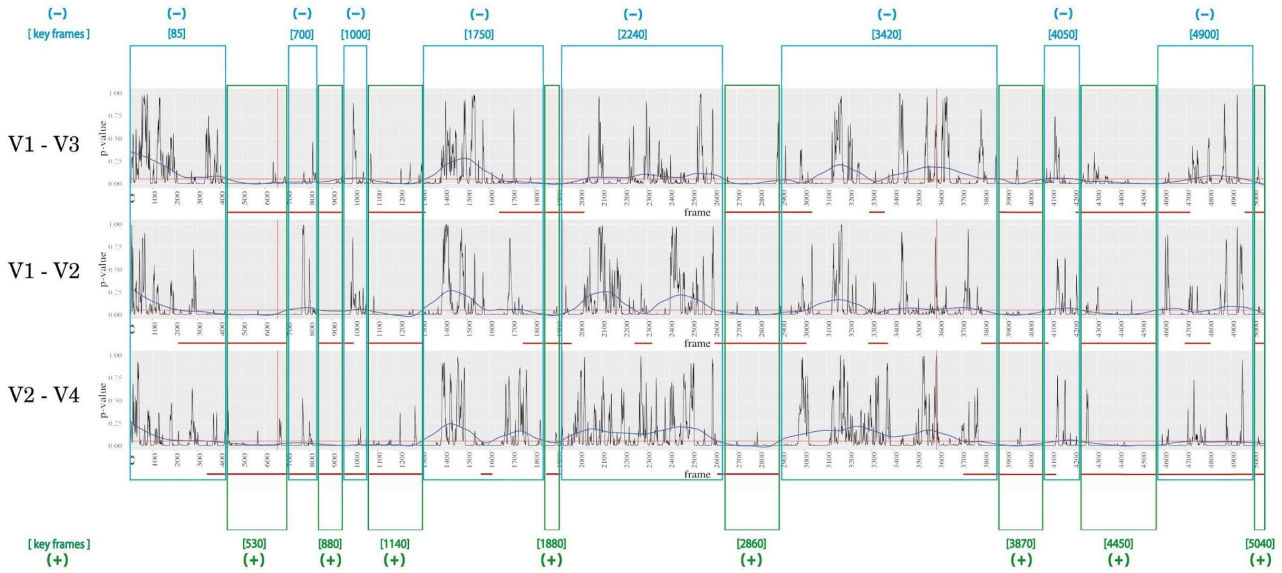


Figure 6. Comparison of PR and NPR videos: episodes statistically similar (-) and different (+).

Following this, with the purpose of globally distinguishing significant and insignificant events, all of the common significant segments were grouped together in the three distributions with a green frame, and the non-significant segments with a red frame. Thus, 16 segments were identified with transversally similar dynamics: 8 segments where significant differences exist between PR and NPR (framed in green) and 8 segments without significant differences between PR and NPR (framed in blue). The segments with significant differences were labeled with a green "(+)", and the segments without significant differences were labeled with a blue "(-)".

5.4.3. Ocular response to cinematic procedures

Comparisons of PR and NPR were made through subsets and keyframes. With the purpose of visually characterizing the 16 episodes, 16 keyframes were selected. The criteria for selecting the frames was that the chosen image corresponded to the segment that represents a moment of the scene where a recognizable "cinematic procedure" or a "cinematic event" exists. Thereby from each one of the 16 keyframes, 4 images were extracted, corresponding to each one of the videos. Thus we analyzed the 16 events from a total of 64 frames (32 NPR frames and 32 PR frames), organized into two groups: eight similarities (+) and eight differences (-) between NPR and PR (Figure 7).

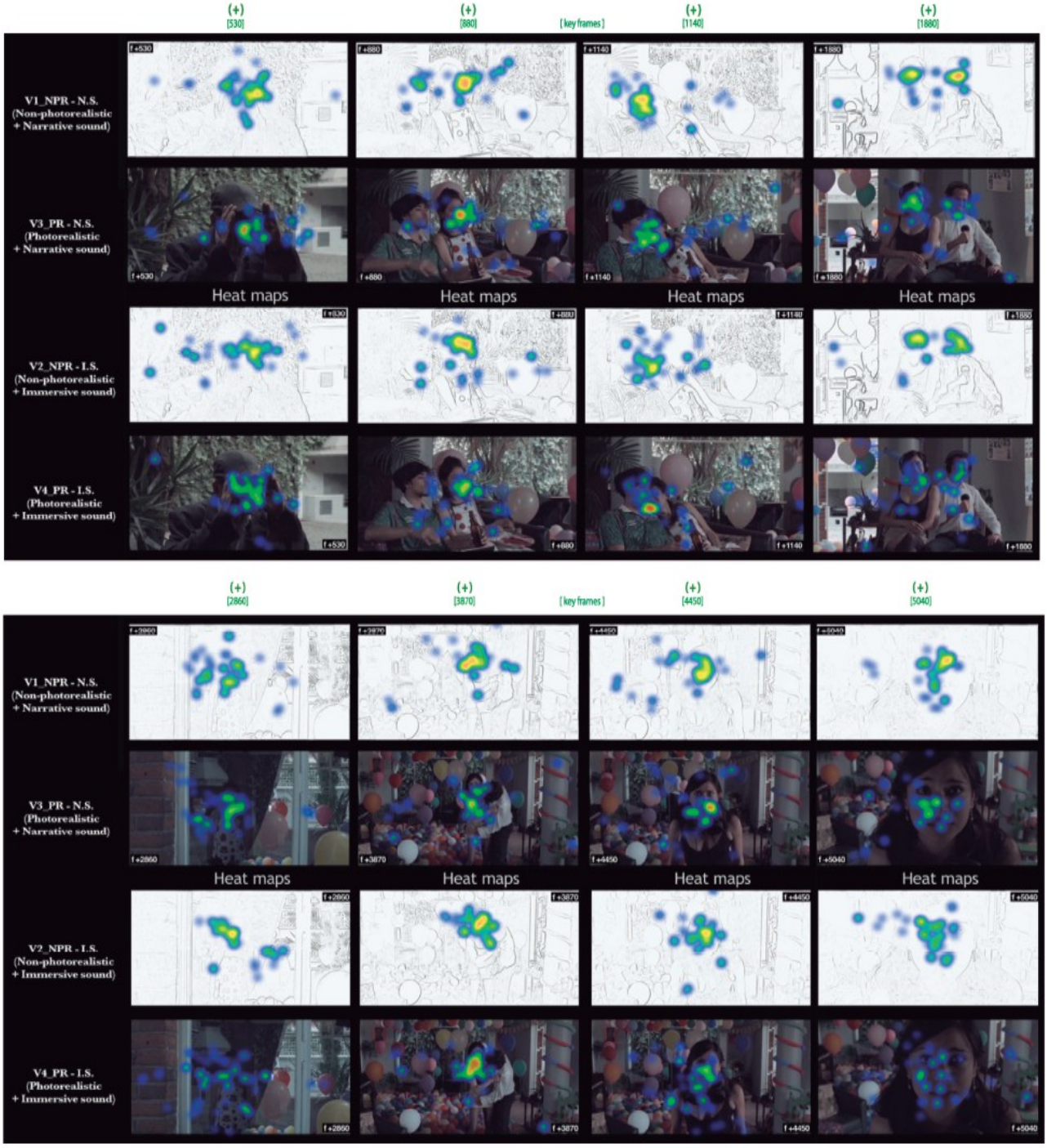


Figure 7. Comparison between NPR and PR: key frames of 8 episodes with significant differences (+). The heat-map visualization is consistent with the idea that there are significant differences between NPR and PR videos: it can be noticed than that there are more concentration of the audience´s gaze in the NPR videos, mostly around faces and characters, than in photorealistic videos.

5.4.3.1. Film editing

The ocular behavior in PR and NPR videos is compared to the editing cuts, to establish common patterns of behavior in relation to these cuts. The film has four scenes. Three edition cuts between scenes were identified. In the first seconds after each editing cut, the gazes took less than a second to switch from an established object of interest —in the previous scene—, to a new object of interest in the next scene (see Figure 8). These comparisons between PR and NPR indicate a similar adaptive response to scene change after each edition cut.

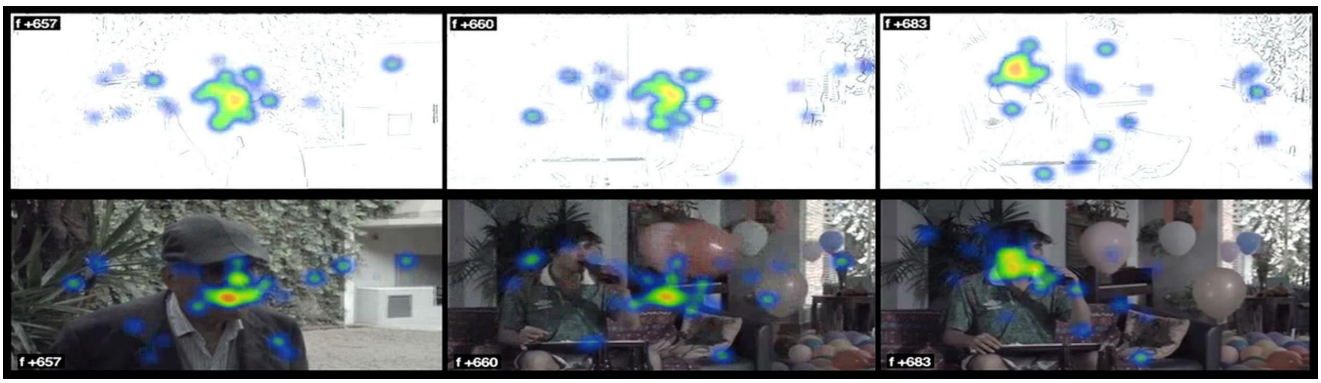


Figure 8. Eye tracking results for film editing.

5.4.3.2. Camera movements

The ocular response to camera movements is compared in PR and NPR modes to establish if there are common patterns. The camera movements present in the film were identified: movements on the Z-axis (dolly In/Out) and movements on the X-axis (traveling Left/Right). The ocular behavior in PR and NPR mode is compared since the beginning of the end of each camera movement. The results associated with the lateral axis (traveling) are clearer than on the depth axis (dolly). The lateral traveling movements are correlated with a motion of gaze in the same direction, immediately after the beginning of the camera movement. The gaze appears to switch from an object of interest, visible before the beginning of the camera movement, to another object of interest that appears when the camera movement discovers a new space of the scene. This correlation is similarly reproduced in movements from left-to-right and in right-to-left camera movements (see Figure 9).

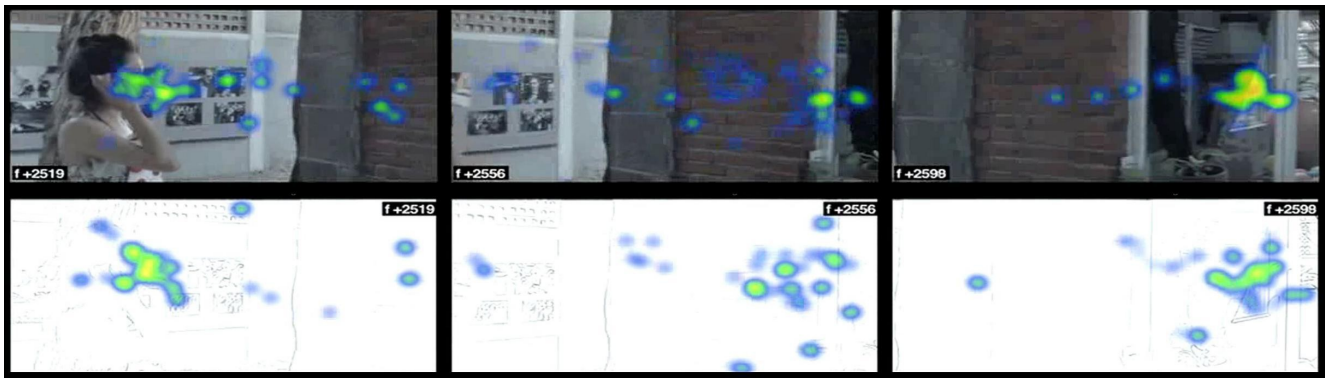


Figure 9. Ocular response to camera movements

Eye tracking results for camera movements.Concerning the movements in the Z-axis (dolly), due to the fact of not using methods of quantification of fixation density in objects of interest, it was not possible to establish a clear correlation between camera movement in the Z-axis and eventual variations in fixation density in the scenic depth. Looking for possible correlations between the cinematic events and the ocular behavior, from the 16 key-frames, representative of the 16 segments, we qualitatively identified 16 cinematic events. Due to the fact that several of these events were repeated, we reduced this number to a total of 9 modalities of cinematic events: 5 types of cinematic events in the episodes with no statistically significant differences between PR and NPR, and 4 types de cinematic events in the episodes with statistically significant differences between PR and NPR.

Due to the fact that we do not use a statistical method to delimit and calculate the density of fixations around objects of interest, we apply a qualitative description of the ocular data in terms of two modalities of gaze concentration: high attentional synchrony and low attention synchrony (see Table 3). This procedure allowed us to compare the ocular perception of PR and NPR videos by an evaluation of which cinematic events seem to be correlated with higher and lower concentration of gaze.

EPISODES WITH NO SIGNIFICATIVE DIFFERENCES (-) BETWEEN PR AND NPR		
CINEMATIC EVENTS	HIGH ATTENTIONAL SYNCHRONY	LOW ATTENTIONAL SYNCHRONY
1. Static scene without characters	General exploration of the scene	General exploration of the scene.
2. Appearance of a character in the scene	Fast concentration of focuses on the character who enters the scene.	Exploration of the depth.
3. Physical contact between characters	Concentration of gaze on the zone of physical contact.	Exploration of depth behind the characters.
4. Characters speaking	Concentration of gaze on the main character.	Exploration of depth behind the characters.
5. Non-speaking characters	Concentration of gaze on characters and spatial surroundings.	Exploration of depth behind the characters.
EPISODES WITH SIGNIFICATIVE DIFFERENCES (+) BETWEEN PR AND NPR		
CINEMATIC EVENTS	HIGH ATTENTIONAL SYNCHRONY	LOW ATTENTIONAL SYNCHRONY
1. Scene with non-speaking characters	PR: Higher gaze dispersion in exploration of the scene.	NPR: Lower gaze dispersion in exploration of the scene.
2. Physical contact between characters	PR: Higher gaze concentration in areas of physical contact.	NPR: Lower gaze concentration in areas of physical contact.
3. Corporal movement on the X axis	PR: Higher gaze dispersion in exploration of the scene.	NPR: Lower gaze dispersion in exploration of the scene.
4. Body movement on the Z axis from the bottom to the front	-PR: Higher gaze concentration on areas of the body. -NPR: Higher gaze concentration on facial zones and areas of physical contact. -NPR: Higher gaze concentration on faces than on the background. -NPR: Higher gaze concentration on the eyes and the center of the face.	-NPR: Lower gaze concentration of areas of the body. -PR: Lower gaze concentration on facial zones and areas of physical contact. -PR: Lower gaze concentration on the face than on the background. -PR: Lower concentration on the eyes and the center of the face.

Table 3. Comparison between NPR and PR in terms of attentional synchrony.

5.5. Discussion

In general terms, the NPR is associated globally with a lower visual entropy, that is, a greater gaze concentration on fewer areas of interest. Qualitatively we may observe two correlations: the NPR is associated with a greater concentration in the faces and a lower exploration of the scenic depth. This result is consistent with previous studies on the perception of NPR. However, when considering the fluctuations of entropy throughout film temporality, we find an opposition between a photorealism cognitively more monotonous than non-photorealism videos. This leads us to consider the visual abstraction of NPR as a modality of cognitive complexity that can not be explained only by a smaller number of salient areas. We will associate the visual abstraction of the NPR with the idea of cognitive uncertainty, but we will build this concept combining neurobiological and sociological perspectives.

The greater density of fixations on the faces could be explained because the NPR generates greater visual uncertainty: the uncertainty generates anxiety, understood as the need to focus on cognitive familiar forms to get out of the strangeness of an abstract environment [GN13] [RC15]. We look for familiar images that give us information to interpret the scenes [WS08] [BM09]. The faces are the sources of the voices and at the same time an indication of the emotional state of the scene [SCH09] [CN08]. In this sense, the faces of characters, rather than the bodies, condense the greatest source of narration in a film (information and emotion). Faces are the key to interpreting the global uncertainty of an abstract or visually strange scene. That is, we do not look at faces because there is nothing else to look at; we look at faces to interpret the whole scene based on our memory. We look at faces because they are naturally “prominent”, that is, a natural magnet for the gaze. The evidence suggests that a biologically innate dimension exists in this behavior [LL20].

From another point of view, it is possible to differentiate two focalization processes detected in the stereoscopic film: the abstract image has the tendency to concentrate on the face, and the photorealistic image suggests a greater tendency to explore the scene. It is feasible to interpret this result through the place that the face has occupied in the cinema. Jacques Aumont describes a historical evolution of the face in Western culture [Aum92], pointing out that its symbolic importance appears only after the Renaissance, and later, with the birth of classic narrative cinema. Then, we can conclude that the cultural importance of the face is a relatively recent phenomenon. This cultural history of cinema seems to be coherent with the biological notion that the face has an innate importance.

Our point of view will be to try to combine both notions to generate the following interpretative hypothesis for the results. Our results seems to indicate that, in a photorealistic style, the face is not always and necessarily the most salient object, thus the gaze may explore other objets of interest in the scene. On the other hand, when faced with an abstract scene, they try to establish an interpretative thread that does not make them lose themselves in the temporal development of the film. In this sense, they use the cultural memory of the so-called "hollywood style of filmmaking" [LLMS15], and choose to focus their gaze on the face of the characters, as a strategy which tends to simplify the interpretive plurality of an abstract scene. The NPR leads us to search in the faces the keys to infer expected behaviors and interpret the global scene. This seems to apply especially in the context of audiovisual content whose coherence is guaranteed by the narrative centrality of the characters. This interpretation is consistent with the neurobiological conception that establishes that the subject perceives from his own cultural memory. However, the real significance of the face as an inferential mechanism should be confirmed in future investigations, with a sample of subjects from other ages of life and from different visual cultures. In this sense, we understand these results as a heuristic contribution to understand working correlations between visual entropy and narrative focus in abstract audiovisual environments.

5.6. Conclusive remarks

We concluded that a NPR filter associated with a high level of abstraction could trigger an adaptive ocular response that may be described as perceptual anxiety: a lower exploration of the scene and a greater focus on the more expressive elements of the scene: bodies bodies and faces of the characters. This ocular behavior, triggered by NPR, could benefit cognitive immersion, especially in cinematographic genres (fantasy, horror, suspense, science fiction) that use visual abstraction and cognitive uncertainty as an aesthetic effect to favor greater involvement of viewers in their narrative movie.

6. Tracking gaze behavior: pupillary reactivity to NPR in 3D cinema

6.1. Introduction

One of the aims of this thesis 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.

In this chapter, 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.

6.2. Related research

In a study about immersion on film viewing [VTM10] 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 [FMCOS17] 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. In chapter 5, we observe the attentional effects of NPR considering the dynamic behavior of the ocular fixation [FGMDYMMS18]. This led us to review similar studies in the same field. In a previous study (see Chapter 5), we found interesting evidence, not necessarily generalizable but highly suggestive: the NPR may capture the viewer's attention with moderate abstraction [PR16] and with the presence of faces [FGMDYMMS18]. 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 [PR16].

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” [HHD04], 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 [HBL18].

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

[HBL18]. 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 [BMEL08]. 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 [MGVP04]. In situations of greater analytical and cognitive demands, it seems to be associated with greater pupillary dilation [MGVP04]. 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.

6.3. Methods

Combining quantitative and qualitative techniques, we analyze the data to establish if there are possible correlations between the audiovisual content and the observed ocular reactivity. The detailed methods are presented in the Chapter 3 “Methods”.

6.4. Results

6.4.1. 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.

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

We established comparisons between the videos regarding their entropy, in order to identify if possible correlations exist between the visual style (NPR, PR) and the 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 10.

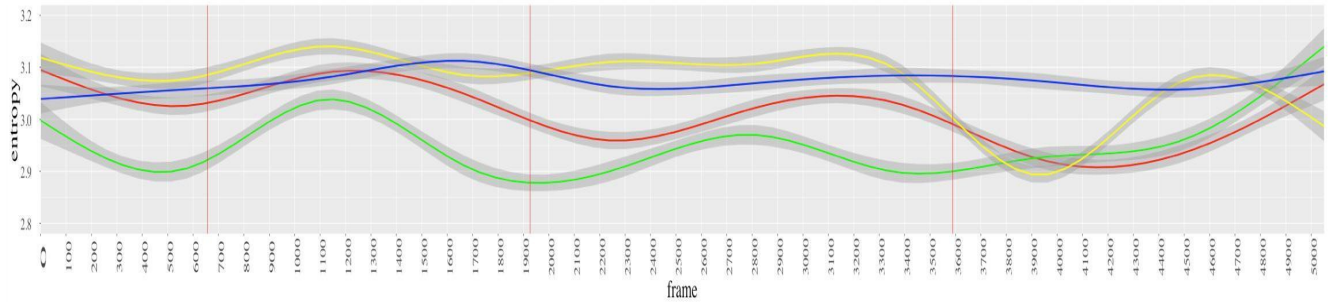


Figure 10. Attentional entropy: ocular response to the 4 short films in terms of visual entropy. NPR-IS [green curve], NPR-NS [red curve], PR-NS [blue curve] and PR-IS [yellow curve].

6.4.3. 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 11.

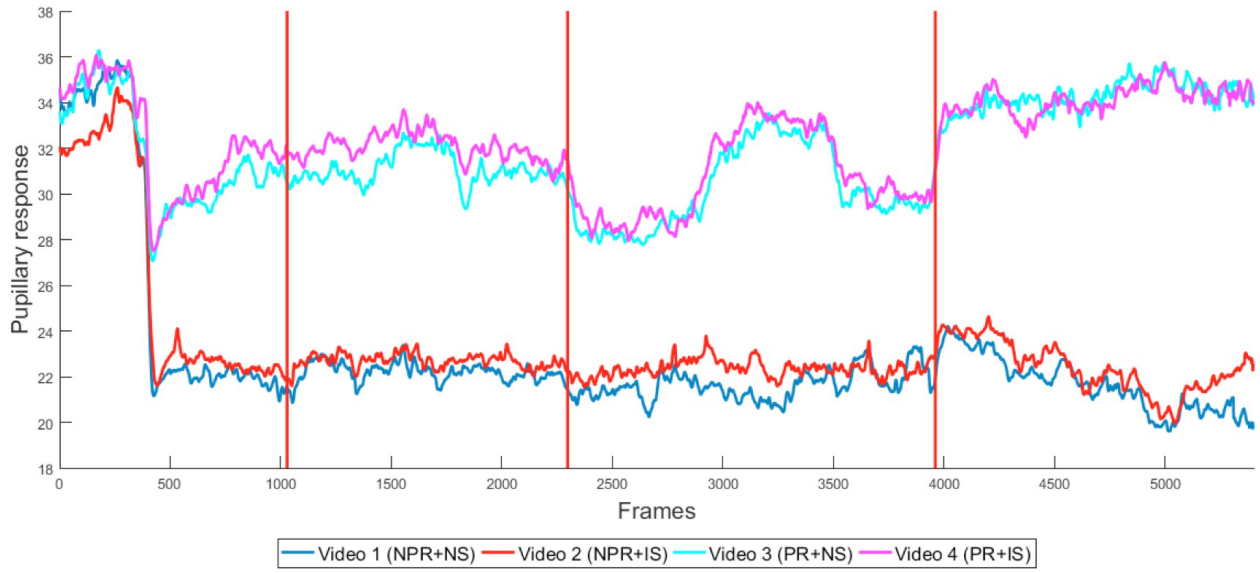


Figure 11. Average of the pupil response of 27 subjects compared to the presentation of the 4 short films. NPR-NS [dark blue curve], NPR-IS [red curve], PR-NS [blue curve] and PR-IS [pink curve].

6.4.4. 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 12.

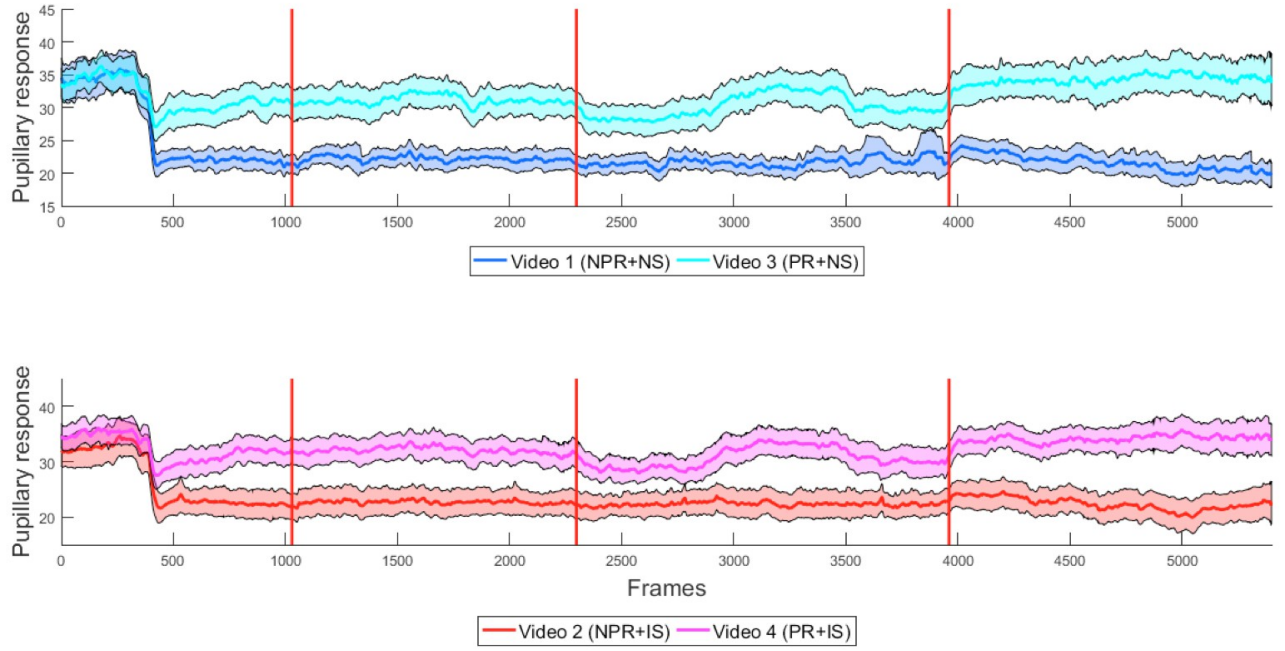


Figure 12. (A) Average of the pupil response for the presentation of the short film with sound type 1 in both NPR and PR. (B) Average of the pupil response for the presentation of the short film with sound type 2 in both NPR and PR. The envelope to the average has the same characteristics as in A).

6.4.5. 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).

6.4.6. 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 13.

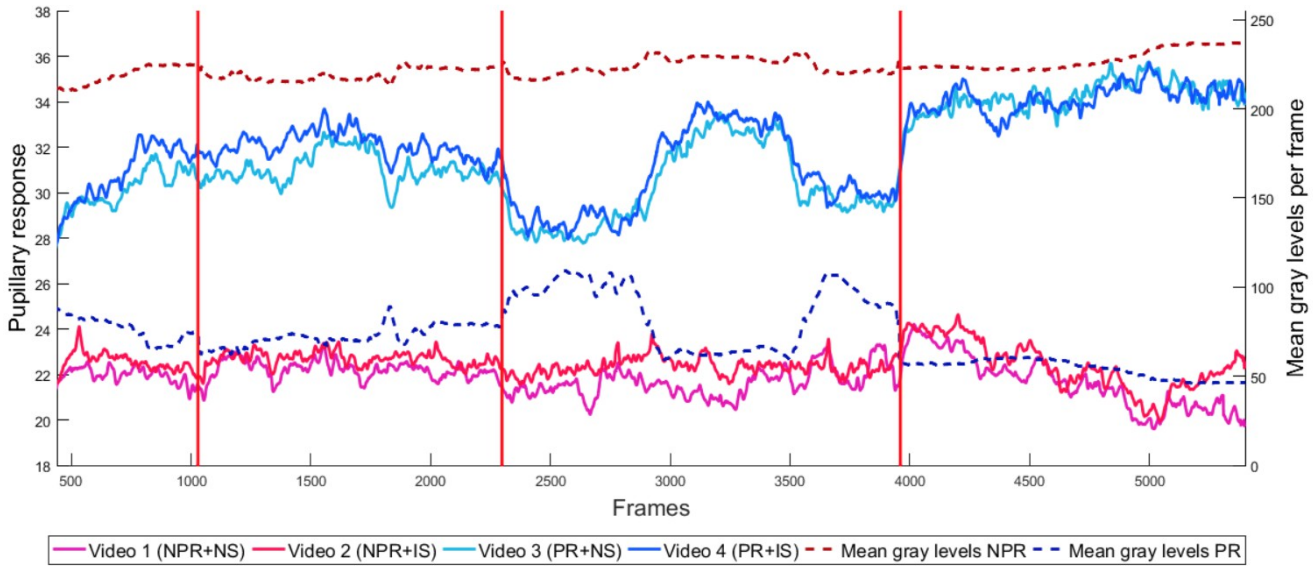


Figure 13. Pupillary reactivity to the intensity of light and post- production in NPR and PR videos.

6.5. Discussion

Our previous research [FGMDYMMS18], described in Chapter 5, 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 [FGMDYMMS18], 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 [Del12] 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 [Del12], 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 significant pupillary fluctuations observed in PR videos, the NPR filter used seems to produce a normalization effect in 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 emphasized the voice of the characters, seems 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 [USS11]: 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.

6.6. Conclusive remarks

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.

7.1. Introduction

In this chapter, we present a research about ocular behavior during immersion in virtual reality. To enhance our comprehension of human behavior in immersive conditions, we will focus on vision and its reactivity to non-photorealistic rendering and cinematic procedures. Using physiological data collected in real-time with an eye tracker, we compare attentional and pupillary response in order to describe the visual immersion experienced by the user. This proposal could be useful to identify physical sickness produced by perceptual mismatches and to evaluate the cognitive efficiency of virtual reality experiences in different fields of applications, giving feedback to content producers about user's ocular behavior.

The field of immersive experiences could be defined as a transdisciplinary design area and development of various products that integrate technology and creativity, which contribute to the creation of experiences in various sectors. When talking about immersive technologies we refer to virtual, augmented and mixed reality, while video games are related to the concept of interaction. However, both technologies can be combined into immersive and interactive experiences. Numerous film and audiovisual festivals in the world have created sections to disseminate audiovisual content produced with these technologies, however, after reviewing the editorial definitions of these exhibition instances, we realized that there is no global consensus on what truly defines this new generation of contents. Some define virtuality as an extension of existing means, importing narrative or interactive operations from the cinema or video games.

Lastly, the point of view that seems most open to us is the definition of virtual from the user's point of view. In order to clarify the contribution of our research we will begin by defining what we understand by immersion and how it relates to the processes of human perception.

7.2. Related research

7.2.1. What do we understand by immersion?

In media studies, this concept is known as “immersion”. This word has its origin in the experience of “diving into” an environment; that is to say, to enter and participate bodily in a new scenario, in a specific space-time.

In a broader sense, this implies disregarding real-world references. And along with that, it allows to address sensations of reality that come from the new environment, created artificially by some means that seek —or perhaps not— the emulation of reality as we perceive it. For experts, the concept of “immersion” has not reconciled a unique and invariable definition. And likewise, the concept has been correlated with other terms, has acted as a synonym for *presence* and has been known under other names (e.g. “immersive presence”) [Ded09].

The concept of “immersion” comes from the idea of experiencing full submergence in a given environment. In a field familiar with this term, such as virtual reality, it relates to the “perception” of feeling physically present in a simulated environment. Virtual reality conveniently provides input devices, with a degree of perception of self-movement, which allows the senses to be isolated enough to feel transported to another place.

Immersion materializes in unique experiences for the user, such as reading a book, comics, watching an audiovisual or musical show. Immersion is studied in different fields and a virtual reality experience may not be immersive. However, for the purposes of our field, immersion will be understood as the objective level of sensory fidelity provided by a virtual reality system [Sla03].

That is, the ability of a simulation to approximate a real experience. For this, the sensations of the real world must be eliminated as much as possible, and replaced by the sensory experiences corresponding to those of the built virtual environment [Mes05]. “Presence”, often correlated and at the same time confused with “immersion”, is understood as the subjective psychological response of the user to the simulation [Sla03]. It is a cognitive and perceptual consequence of immersion, and if wanted, possibly the last phase of this, after the stages of commitment and absorption of the experience [BC04]. In simple terms, it corresponds to a hoax; the subjective belief that one is in a certain place, even when physically located in another [WS98] [PCSRVC00].

Hence, we cannot speak of “immersion”, and consequently, of immersive technologies without referring to the sensory and cognitive impact, ultimately, the bodily effect produced by these digital deceptions. Until now, the gradual realism of the experience seems one of the necessary conditions for the production of previous results. However, we consider it appropriate to specify the “sense of realism”. On the one hand, we have “perceptual realism”, which is what we vaguely understand as “realism” or “photorealism”: “how well the environment looks and sounds like the real world” [McM03]. And, on the other hand, social realism, the plausibility of a reality experience, regardless of the style of representation (e.g. the use of NPR in immersive environments) [GW02].

This condition of reality must be coupled with the technological means used. Virtual reality (VR) is perhaps the computer technology that, in essence, allows through head-mounted displays (HDMs), which cover the user's entire field of vision, the multimedia simulation of a fully immersive scenario. Today, it seems to us an almost natural system of interaction and control of the “immersive” experience, admitting a degree of “perception” of one's own movement, while at the same time, allowing the senses to be isolated enough to feel transported to another place (often confused with reality).

The use of VR, however, has been associated with certain discomforts; being unknown that the possibility that these annoying body manifestations were actually symptoms of increased immersion and/or presence in the face of experience [DD18]. Despite the enormous technological progress, problems of discomfort have been reported in users. They have been studied extensively by the academy, generating standardized metrics and procedures to assess these discomforts. And with this, improve the experiences to minimize these effects, which are mostly related to the depth of fields, experience control, color tone within the experience. This occurs in artistic-cultural content (films, games, educational experiences) as well as in those destined for industrial training of professional competencies.

VR experiences still produce physical discomfort because all the key elements of the associated physiological phenomena have not yet been deciphered, at least in the context of content producers. In other words, creators and engineers produce content without having methods or tools that allow them to evaluate the physiological impacts of their products. Many resources can be invested to build immersive experiences that produce collateral effects, without knowing well how to measure them to avoid them, except going through merely intuitive trial and error processes, without scientific basis.

Our research seeks to advance in the understanding of virtuality from user experience, that is, how people perceive their own experience of virtuality and, therefore, how the discomforts associated with current forms of immersion are experienced. Hence it is not possible to talk about immersive technologies without taking charge of the biophysical dimension of their effects on the user, because most of the experiences that are generated, in startups or audiovisual companies, are carried out without studying the sensory impact of its contents.

7.2.2. Biophysical dimension of immersive experience

3D cinema and VR share a series of similarities emphasizing first of all that both require binocular optical devices. They are also technologies that have expanded the areas in which they are used, such as cinema and museum exhibitions, and there is a massification to the point that they are beginning to enter, incipiently, into homes.

Both technologies use perceptual resources that stimulate multi-sensorially, which implies the use of more than one sense to increase the realism and immersion of the user. In most cases it is bimodal, that is, it occupies two of the five senses that in this case are visual and sound.

According to studies [KKO15] [HVC12], vision is the strongest sense therefore the attention and position of objects in space are determined by the user's vision. For this reason, sound becomes a complement of the image. The above is known as the ventriloquist effect or visual capture effect where there is an impression that the sound originates from the object, a phenomenon similar to the performance of the ventriloquist and his doll.

When perceiving images and sounds from a virtual world it is possible that physical discomforts will be felt by the user, similar to those known as motion sickness, where the person may feel nauseous, dizzy, headaches or eye fatigue [TMF19]. There is a precision in the name which depends on whether it is 3D, where it is known as “3D vision symptom” and in “VR visually induced motion sickness” (VIMS). These technologies simulate realities that engage perception, and this is where “cue conflict” is generated, known as one of the most accepted theories, which refers to the specific information received by the brain and the motor system from sensory stimuli. That is, “cue” is the way the body estimates its movement from perceptions. Therefore, if there is a disagreement between the stimuli and the body, the discomfort that causes the side effects mentioned above appears.

Other elements associated with the discomfort could be related to different factors of each individual where age, concentration, previous experience, gender, or diseases (physical or psychological) can influence. In addition, there are factors directly related to technology that could influence the user's physical feeling such as calibration, contrast, field of view, color, binocular vision, among others [Kol95].

Given that the phenomenon of sickness in 3D or VR is not new, the physical discomfort tests have expanded to the point that it began as a subjective questionnaire that is still applied and today these studies count with medical devices known as the objective examinations. The objective tests include monitoring of breathing, heartbeat, eye blinking, body stability tests or the central nervous system from an electroencephalogram (EEG) [DSS18].

The "simulator sickness questionnaire" (SSQ) consists of a questionnaire that distinguishes 16 symptoms in nausea, discomfort, disorientation and ocular categories, from a 4-point scale where 0 is nothing and 4 is severe. It is commonly used before and after 3D or VR technology to ensure that there are no pre-immersion symptoms [BRR07].

In a study [29] they examined whether the cause of physical discomfort could be induced by exposure time and if the symptoms of VIMS worsened or changed with peripheral vision or the visual monitoring of an object. They used subjective and objective measurement methods to analyze users during the study. Among these they included the SSQ, body stability and near-infrared spectroscopy that measures without invading the body, the reactions of different activities in the cerebral cortex to the subjects in their natural state. They were able to conclude that sensory areas, such as vision, could be activated when using 3D.

3D cinema and VR share the problem of causing physical discomfort during the exhibition, affecting users and immersion experience. The difficulty is that when a user feels discomfort it is possible that there will be a reduction of immersion, directly influencing the user experience in a negative manner. Hence in these secondary symptoms, where the visual sense predominates, it can remove the person from their connection with photorealistic viewing, taking into account that longer exposure time generally increases 3D vision symptom and VIMS.

Each of these discomforts increases or decreases according to the user's susceptibility, as well as the exposure time or the delay between physical movement and its virtual action [Kol95] [Bar04]. However, the emphasis on the influence of visual representation technique (e.g. non-photorealistic rendering) has not been properly placed.

7.2.3. Visual dimension of an immersive experience: photorealism and immersion

In computer generated images (CGI), we find two great methods of expression. On the one hand, "photorealism" (PR) and on the other, "non-photorealism*"* (NPR). The "photorealistic rendering" aims at the optical representation of the world, that is, the imitation of the physical principles of photographic capture, both on volumes, textures, shadows, as effects of color and lighting [Gen10].

In contrast to the PR, "non-photorealistic rendering" gathers and designates a wide variety of creative and stylistic filters of digital processing, and is validated by it [KRHC15]. The NPR before the PR is considered more expressive and beautiful, as well as more efficient in the transmission of information and convincing in performance [HVC12]. Therefore, it constantly has applications depending on the transmitted aesthetics and the ability of scientific visualization that this expressive range gives us [CD14].

This method (NPR) allows you to regulate the information in the visual file (image or video), either by subtracting attributes or by providing new features (e.g., the painting styles inspired by Van Gogh, Renoir and Dufy) [KM06]. Not satisfied with that, within the range of NPR, we find endless possibilities of resulting styles; as well as efforts to classify them according to the design characteristics and behavior of each technique, considering even the interaction with the user [KCWI13].

As we said, NPR has the ability to replicate an artistic style by abstracting the properties of the corresponding image or video [Win06]. Through this process there is not only a modification in style and content of interest, but also the attention and the way it is looked at. In other words, it modulates the perception of images by users, already recorded through questionnaires [HMHLS03] and based on eye tracking data [SD04].

In fact, in previous studies, we have realized that the high level of abstraction could trigger in users: “a lower exploration of the scene and a greater focus on the more expressive elements of the scene bodies and faces of the characters” [FGMDYMMS18], see chapter 5 “Attentional reactivity to NPR in 3D cinema”.

NPR, through the use of diverse styles, refers to different and significant user reactions. On the one hand, it is known that emotions are stronger in the face of photorealistic images. Instead, compared to stylized images, we have a more neutral state. Probably, this decrease in response in the transition from PR to NPR is not due to the loss of information itself, but it would be presumed that NPR eventually confuses and distracts users. Despite this, photo-abstractive processing, a visual format that reconciles the best of PR and NPR, preserves a very good emotional response [MML11].

Considering that “immersion” is a process that, under ideal conditions, should occur without physical discomfort, we will develop an analytical method that allows us to evaluate the importance of visuals in the origin of such discomfort. Our work plan is as follows: we will take the method used in the study of immersion in 3D cinema and adapt it to the study of virtual immersion, in order to evaluate how non-photorealism can modify the perception of content, and consequently, identify possible ways of solution to the discomforts of immersion.

Our hypothesis is simple: inferences about ocular behavior in 3D cinema, obtained through the use of eye-tracking, can be applicable to virtual reality in order to describe the ocular behavior, and to describe the immersion process of each user. The justification for this hypothesis is found in the understanding of human perception as an enactive process, that is, the human response to experiences comes mainly from their socio-historical experience. This implies that each user can combine different types of immersion, and what matters to us, rather than looking for universal patterns, is to be able to correlate the contents and visuals in order to identify the cognitive importance of the various objects of visual interest existing in the immersive content.

7.3. Methods

Combining quantitative and qualitative techniques, we analyze the data to establish if there are possible correlations between the audiovisual content and the observed ocular reactivity. The detailed methods are presented in the Chapter 3 “Methods”.

7.4. Results

7.4.1. Attentional behavior: perceptual anxiety in non-photorealistic scenes

In general terms, NPR is associated globally with a lower visual entropy, that is, a greater gaze concentration on fewer areas of interest. Qualitatively we may observe two correlations: NPR is associated with a greater concentration in the faces and a lower exploration of the scenic depth.

Facing a non-photorealistic scene, characterized by a correlation between visual abstraction and cognitive uncertainty, the gaze seems to react under the effect of a perceptive anxiety: an ocular behavior that can be described as a limited exploration of the scene and a greater focus on the most expressive elements of the scene, such as bodies and faces of the characters. This behavior appears to us as an adaptive response to a visual environment endowed with abstraction and uncertainty. This perceptive anxiety seems to spontaneously assign greater sensorimotor resources to the analysis of the character's behavior, transforming his face into the main key to analyze the information and the emotional tone of the global scene.

NPR leads us to search in the faces the keys to infer expected behaviors and interpret the global scene. This seems to apply especially in contents whose coherence is guaranteed by the narrative centrality of the characters. But it could also be a good clue to explore the attentional behavior in virtual reality contents with 360 degree scenes, where the user is constantly invited to move their gaze to explore the scenic space.

7.4.2. Pupillary reactivity: emotional immersion vs. intellective immersion

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 significant pupillary fluctuations observed in PR videos, the NPR filter used seems to produce a normalization effect of 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. However, we found a result that changes everything, because it suggest the appearance of a possible emotional response mixed with the pupillary reactivity to light.

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 performed a doubly significant action: at the same time that she developed a verbal discourse with an intensifying facial expression while she approached the camera (see Figure 14). During her body movement towards the camera, the stereoscopic effect produces the illusion that her 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. This final scene is more visually complex than the previous ones and 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.

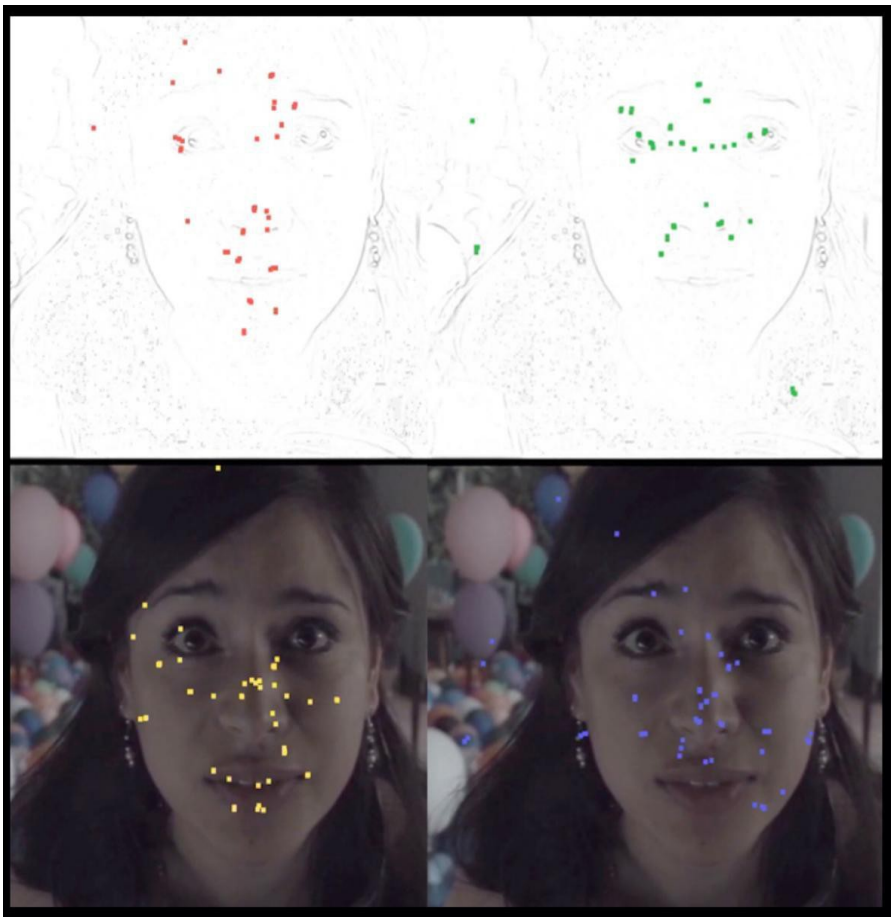


Figure 14. Final scene of the film analyzed with eye-tracking data. the colors of the points were differentiated only to be able to compare the spatial distribution of the spectators' gaze in the image space.

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 understand cognitive complexity as a central factor to understand immersion. The case 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.

Then, everything is played between the space of the perceiving subjects and the type of visual style proposed. 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.

7.5. Discussion

The results obtained indicate that ocular behavior, described thanks to the use of eye-tracking, can be a useful source of information to describe immersion. When describing the evolution of the vision in relation to the temporal flow of the content, it will be possible to interpret the reactions of the users to the fluctuations of the characters and the scene.

The concept of perceptual anxiety could be useful in the context of 360-degree scenes, since it would identify the user's need to obtain the information necessary to interpret the meaning of each scene. Unlike the cinematographic montage, prepared before viewing, the head and body movements made by VR users constitute the new form of real-time editing. If we are able to characterize the perceptual functions of these vision movements, we can analyze the narrative and cognitive efficiency of a virtual reality content based on the user's eye behavior.

Combining the analysis of visual entropy (dispersion of the vision in the scene) with the emotional response, we will be able to have experimental evidence, subsequently verifiable through interviews, to discern how intellectual and emotional moments could alternate. There is no single type of immersion: each user travels the scenic space according to their own interests, but if we discover recurring patterns in the ways of seeing, we will have advanced another step towards the understanding of immersion in VR.

These inferences about vision were made possible by the experimental use of NPR. If we use these findings to assess the importance of visual factors in the occurrence of discomfort during virtual immersion, we can replicate the same method: evaluate whether the same virtual reality experience, operating with different visual styles, generates physical discomfort at the same times. The visual abstraction of the NPR seems to elicit different responses from the viewer of 3D cinema. If this is replicated in virtual reality, it would mean that post production, together with ocular analysis, could reduce or modify the negative effects of virtual immersion.

7.6. Conclusive remarks

The comparative analysis of ocular behavior under different visual styles opens up new perspectives for the understanding of the concept of immersion. Attention and emotional vision markers can be used to implement an empirical method, based on eye-tracking data, to study immersion in virtual reality. In order to deepen this line of research, it will be necessary to adapt the methods of data analysis, for example, by incorporating quantitative methods of ocular data analysis to identify objects of visual interest in the three-dimensional space of immersive experiences. If we can correlate the way we see with what interests us in a scene, we can describe our immersion experience with greater precision.

8.1. Introduction

The appearance of audiovisual narratives that are based on the use of interactive media and immersion require the development of visual methods for its study. It is fundamental to understand how virtual reality and interactivity change the cinematic language, transforming its expressive mechanisms, and the spectator/user experience. This represents a great challenge for audiovisual production and scientific research, due to negative physical effects such as dizziness and other nuisances that can be generated by some of the immersive content. We have developed a transdisciplinary approach destined to identify narrative and interactive procedures that enable the reduction of collateral effects of the filmic immersion of virtual reality. This method is based on studies on cinematic visual perception, and it incorporates a collaborative audiovisual production process, the design of implicit interactions between the content and the user, and the evaluation of real-time physical and emotional answers during the immersive experience.

8.2. A transdisciplinary approach to cinematic and technological research

The Virtual Reality Laboratory [VR-LAB] of the University of Chile is a nucleus of experimental scientific research and an artistic audiovisual creation dedicated to the study of immersive audiovisual languages and emerging interactive digital technologies. We are interested in studying the immersion of individuals, groups and collectives, in virtual worlds and environments that allow spatial navigation and digital communication, with a transdisciplinary academic approach.

The perception of the virtual as a complex phenomenon implies a “constructed corporality”. The virtual user inhabits a new space, perceives new limits, new distances, their vision shifts, their attention changes; in summary, their perception is modified. However, this perception, because of its mediation role between the inside and the outside, in turn transforms the constructed space. From where, then, should one begin to approach the complexity of the virtual? The mode of disciplinary knowledge production allows, without a doubt, the outline of positions, but the juxtaposition is not enough to account for the pace of innovative technology that underlies and gives feedback to the “virtual reality” object. The search then leads to the construction of a dynamic map, an analytic puzzle that manages to transcend the

speedy technical progress to become a changing perception, capable of apprehending the complexity of its objects.

Transdisciplinary, understood as transcendence of disciplinary, academic and epistemic limits [TOLLC17], appears as a possible route that, although it does not guarantee success, it could provide a new conceptual architecture and a new experimental outlook to account for this new "virtual corporality". "Transdisciplinary research includes academic researchers from different disciplines and nonacademic participants that research real world problems and create new knowledge and theories" [Cro08]. In the construction of our visual method, we added a new level of complexity: representation. In this case, the virtual is applied to an audiovisual object, a spatial representation. The approach implies gathering users and experts to experience and perceive this virtuality. These participants are not only "subjects of study", but their perception also modifies the outline of the objects. The transdisciplinary view comes forth through the integration of artists and scientists, outlining a methodological framework that enables their participation and synergy between their knowledge.

8.3. Immersion in virtual reality and its obstacles

Within scientific literature, the concept of "immersion" proceeds from the idea of submerging in a determined environment. In a field familiar with this term, such as Virtual Reality, "immersion" is related to the perception of feeling physically "present" in a simulated environment. Virtual reality conveniently provides input devices, such as headsets and helmets, means of almost natural interaction and control, with a degree of perception and auto-movements that, at the same time, allow the isolation of the senses to a sufficient degree that one feels transported to another place. Good practices can increase the intensity of this sensation which has a direct relation with the user experience. Immersion is studied in different fields and a virtual reality experience could not necessarily be immersive. Immersion materializes into unique experiences for each user such as, for example, reading a book, comics or watching an audiovisual or musical show. However, for purposes of our study, immersion can be understood as the objective level of sensory fidelity of something experienced and provided by a virtual reality system [KJK17]. In other words, the capacity of a simulation to come close to a real experience. For this, real world sensations must be eliminated as much as possible and be replaced by the corresponding sensory experiences that match those of the constructed virtual atmosphere. While "presence" or "the feeling of presence", often correlated and at the same time confused with "immersion", is understood as

the illusory subjective user response to simulation [SSv16]. It is a cognitive and perceptual consequence of immersion, that corresponds to an illusion: a subjective belief that you are in a certain place, while being physically in another.

According to the above, it is not possible to refer to “immersion”, and in consequence “immersive” technologies without mentioning the sensory and cognitive impact. Ultimately, the corporal effects that are produced by these digital illusions. Looking back at the initial decades of these technologies, the high digital definition, the computer processing speed, the ergonomics of the existing headsets and the globalized access to these kinds of devices, you would think that the virtual era has arrived. However, despite the enormous technological advances, collateral effects are still being reported in virtual reality experiences [SGCDPSGBG20]. Nevertheless, these problems have been studied extensively by academia, generating metrics and standardized procedures in order to evaluate these discomforts [Sto17]. And, through this, improve the experience to minimize these effects, that are mostly related to depth of field, control of the experience and color tones inside the experience.

The inseparable essence between virtuality and corporality has not been resolved, at a global level, even in industrialized countries where these technologies are more consolidated. This is observed in artistic-cultural contents (movies, video games, educational experiences, etc.) as well as experiences destined to industrial job skills training. Virtual reality experiences still produce physical discomfort because all of the associated physiological phenomena have not yet been deciphered, at least in the content production area. In other words, creators and engineers produce contents without having methods or tools that allow them to evaluate the physiological impact of their products. Large amounts of money can be spent to develop immersive experiences that produce side effects, without any knowledge on how to avoid them, unless by going through trial and error processes that practically no national market or company is capable of paying.

From the above, it is given that it is not possible to refer to immersive technology without taking into account the biophysical dimension and its effect on the user, because most of the experiences generated, in audiovisual companies or startups, are completed without studying the procedures.

Therefore, there is a possibility of not impacting the objective group in a positive way, causing rejection towards this technology and a technological breach. A technology center that aims to stimulate the production and distribution of content produced with these technologies will

necessarily have to assume the challenge of introducing, among the same investigative processes, development and prototyping processes, specific scientific knowledge capable of accounting for collateral effects. It will be necessary to identify these collateral effects in order to correct the possible sources of these unwanted disorders and iterate tests until the contents are brought to a level of such comfort that it will be reasonable to move to the production and distribution phases.

As a starting point, we propose to face this challenge of developing a visual method in order to study immersion and interactive audiovisual content in virtual reality: we wish to study the implications of the biophysical interconnexion between content and spectator-users from a transdisciplinary perspective of its problems.

8.3.1. Virtual movement and physical discomfort in virtual reality

As we explained previously in chapter 7, about the biophysical dimension of immersive experience, within the diversities of discomfort that can be produced by immersive content, we consider the following discomforts: “motion sickness”, “simulator sickness” and “cybersickness”, all of which are differentiated from one another with a certain degree of ambiguity. Even though all of these share a large part of the symptomatic profile such as headaches, fatigue, disorientation, dizziness and nausea, it is not possible to refer to the same discomfort that appears due to movement; at most, the last two could qualify as subtypes of motion sickness [MWHN18]. “Motion sickness” refers to the physiological response to the incongruity between the signals perceived visually and the movement perceived by the vestibular system, responsible for balance and control of space [BS16]. “Simulator sickness”, compared to the previous experience, is that which is specifically due to the virtual environment, either through a projection device or a head-mounted display, that allows simulation, exclusively, piloting or handling some type of vehicle [BRC19]. “Cybersickness”, just like simulator sickness, is produced by the interaction with virtual environments, especially with the assistance of interfaces such as headsets or virtual reality helmets, which are not restricted to only movement, but are also stationary [RO16]. And, until recently, understood to be triggered only by visual stimulus, not involving the action of the vestibular system [APA18].

There have been attempts to reduce and relieve user discomfort, but the results are contradictory, due to the multifactorial nature of the discomforts [CKY20]. However, we can highlight the main factors, related to the hardware and the content, that contribute to the manifestation of these symptoms and the main recommendations to mitigate them. According to [PCTRS20], within the causes of the sicknesses that come from virtual environment there is the interaction and locomotion speed, the high rate of acceleration in mobility, the wide field of view, the inadequate simulation of the depth of field, the low level of control, the long exposure time, latency or lag in physical action and its corresponding virtual representation, the lack of a rest frame during the experience, the use of camera rotation, and postural instability. Within the strategies that reduce discomfort, it has been considered, for example, teletransportation techniques and slow motion, the gradual reduction of the field of view through vignettes, the implementation of rotational motion blurs, the periodic disconnection from the experience, and the consideration of rest frames.

Ultimately, the interactivity of virtual reality expands the possibilities of audiovisual language, but it raises new cognitive and corporal “dispositions” of the spectator-user, as well as specific problems that are part of immersive narratives, such as the perception of movement in virtual environments that produce negative effects on the spectator-user. The negative effects associated with virtual reality represent an important obstacle to immersion, that is to say, the sensation that a spectator-user has while in a virtual space, making them believe it has similar characteristics with reality. The illusion of this virtual immersion can be seen, in consequence, limited by cybersickness. This means that one of the main staging elements (the movement of the characters and/or camera), in the context of interactivity with the spectator-user, will be a recurring source of problems as long as the factors causing the discomfort are not resolved or mitigated.

8.4. Experimental application: a visual reality prototype

8.4.1. Congruence and interactions as key factors for virtual immersion

Even though the displayed method has presented, until now, tools to study perception in 2D and 3D movies, along with some extrapolation proposals to the immersive format, it is essential to incorporate two key aspects in order to study immersion in interactive virtual reality: the experimental consistency between two different components of the virtual world proposed as a requirement to forget its artificial character and the incorporation of interactions that change the spectator into an active user, capable of performing actions that modify their environment. Hereafter, we will explain both factors in terms of their contribution to interactive virtual immersion.

8.4.2. The experimental congruence as a factor of immersion

The experimental results about immersion on virtual environments (VEs) are consistent in showing that immersion is far from improving only with passively perceiving good audio-visual stimuli [SvS05]. This happens because we are complex and self-organized dynamic biological systems [Fre00], for which the conscious experience is the result of the history of organism-environment interactions. From this perspective, perceiving is a way of moving (and vice versa) [Var95]. In other words, perceiving is not something that is done exclusively and passively with our senses, we literally perceive with the entire body [HH63]. Furthermore, what we perceive in our subjective "now" involves a creative synthesis between our past experiences and our future expectations [Var99]. From the idea of the action-perception cycle and the temporal integration of "now", the concept of experiential congruence arises. The proposal consists in designing VEs where what is perceived is naturally linked to the movements of the subject, together with integrating, at the same moment in time, what the subject has just experienced with what is expected to happen in the future, given the history of Subject-VEs interactions.

8.4.3. Conventional and non-conventional interactions

The design of implicit interaction mechanisms is largely related to human behavior and its physiological and psychological capacity to learn new skills. In virtual reality environments, users can behave naturally, interacting with objects around them as they would in the real world [LaV20], although each person would do so in a different manner. In these systems, implicit interactions would predominate, that is to say, it is about capturing the will of users in

unpremeditated movements [BJ19]. In contrast, there are the classic interactions that are carried out, for example, through a graphical interface to communicate the intention to perform a certain action. [Ju15] contrasts the difficulty with classic interactions, being insufficient to relate or put oneself in the shoes of the user. In addition, it highlights the search for patterns in the actions of the users to incorporate them into natural movements, some of which having been studied in other disciplines. [Sch00] points out that it is important for the system to know in which context the actions take place (e.g, analyzing body language) and to adapt its responses to it.

8.4.4. Audiovisual creation and experimental research on virtual immersion

Given that interacti86vity appears as the key factor to understand this new modality of documentary narrative [UDWB16] [AA10], we will define two forms of interactions: explicit and implicit. In contrast to explicit interactions, implicit interactions are defined as those forms of interaction between human and computer that do not require conscious action on the part of the user [JL08]. Our method will consist of observing how the user enters the experience, analyzing both forms of interaction.

The spectator-user will be immersed in the documentary representation of confinement spaces. They will be able to explicitly interact with some objects and cause implicit interactions with specific objects or areas of space. Just as audiovisual editing is a key factor in favoring the continuity of attention of the viewer throughout the film [SM13], we will define three forms of user movement to favor the sensations of immersion in the story. Immersion can be understood as the target level of sensory fidelity provided by a virtual reality system [Sla03]. For this, real-world sensations must be removed, as much as possible, and replaced by sensory experiences corresponding to those of the constructed virtual environment [Mes05]. An example of sensory experience is movement. We will define three types of virtual movement: the user will be able to move into a room and teleport through a “portal”, in two ways: using a portal “A” that is physically crossed when moving from one room to another, and a portal “B” that, when touched, produces a direct “jump” between rooms.

Our methodological proposal will allow different types of data to be correlated in order to build a documentary immersion map, that is, a scheme that allows describing the different ways in which a viewer discovers, navigates and interacts with a documentary in immersive virtual reality.

In [FWMS19] it is stated that: 3DS cinema and VR share the problem of causing physical discomfort during exposure, affecting users and the immersive experience, as it would reduce it. And, in order to evaluate how NPR can modify the perception of the contents, and, consequently, identify possible ways of solving the discomforts of immersion, we propose that the method used in the study of immersion in 3D cinema can be adapted to virtual reality. Correlating to this, for the navigation of spaces in stereoscopic virtual reality (VR-3DS), we propose a method with two narrative modalities of immersive VR experiences. With a transdisciplinary approach, we intend to analyze through the virtual interface, that is, the narrative conceived by the multimedia project team and the quality of the user experience observed in real-time.

The contribution of the audiovisual creation to research, in this case, is based on the conviction that the artistic language of cinema, expressed in visual and sonorous attributes, is a fundamental basis to experimentally validate the generated hypothesis from a scientific perspective. The audiovisual research process, development and production of the immersive short film called “Insomnes VR” will be implemented as an iterative process, where we will be able to apply the results of previous research, evaluate if they give the expected results, and in this way, consolidate artistic options thanks to the verification of scientific hypotheses, in a transdisciplinary aspect.

The “Insomnes VR”project emerged during the pandemic, between March and April of 2020. Due to the uncertainty of this global health crisis, we defined an experimental approach to documentary language that incorporates, as a forced footing and as a characteristic feature, the option of not making field records. Due to the restrictions of the sanitary lockdown, that lasts until today, we decided to produce all of our material under the confinement conditions. The reference to the real will be produced from visual, sonorous and verbal testimony records of people who live confined in different conditions. The title “Insomnes” comes from a common condition (insomnia), adhered to mental health such as anxiety and sleeping disorders, reported from people of different ages, who have had to live their mandatory lockdown under stress conditions, sometimes motivated by the loss of their job, family division, senior’s life risk, the interruption of all physical contact, and more. The pandemic “insomnes” are men and women that have seen their quality of life deteriorated due to the accumulation of weeks and months of sleep deprivation. Insomnia is the metaphor of the lack of control in the face of a reality that surpasses us as a society and as humanity, in different degrees according to each country.

Next, we will explain some of the stages of the creative process destined to represent, in virtual reality, the confinement conditions of these “insomnes”, in closed (room) and open (ship) spaces. The creation process is a collaboration between the audiovisual team, the scientific team and people who shared their confinement experiences.

8.5. A virtual reality prototype: a gaze-based experience

8.5.1. The design of immersion for closed spaces

The first step to represent a closed confinement space (see Figure 15) was the modeling of two rooms using the software Blender (https://www.blender.org/), incorporating textures to later export them to the software Unity (https://unity.com/es).

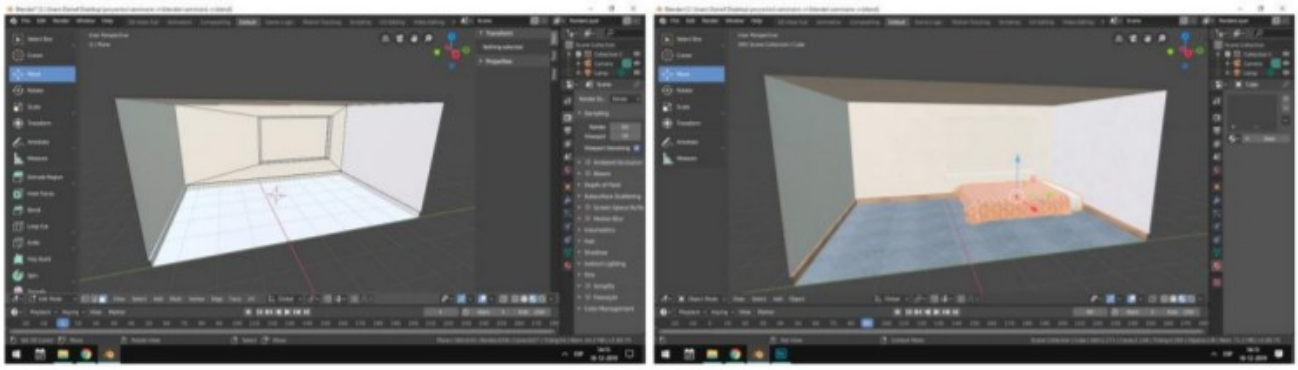


Figure 15. Room modeling

Once the rooms were created, windows were built to create exteriors with their own dynamics (see Figure 16), that contributed to the feeling of confinement, which in some cases even allowed interactions (such as scaring a cat on the roof) that increased the impression of realism.

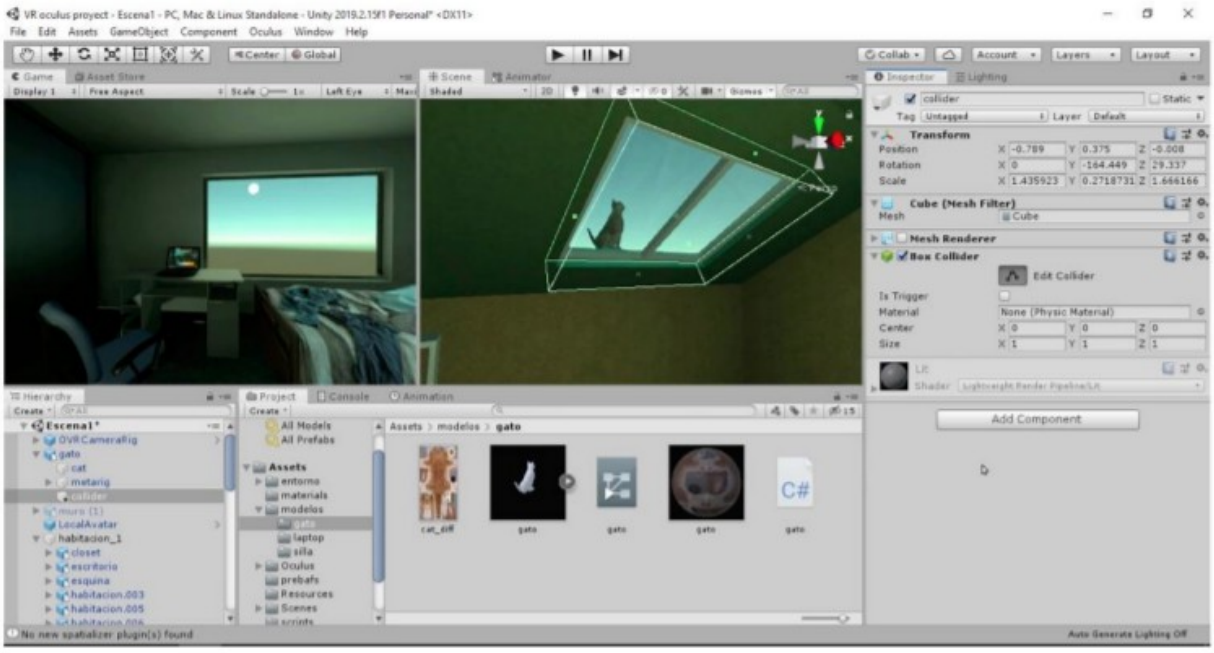


Figure 16. Creation of interactive dynamic windows and exteriors.

In order to simulate the possibility of moving between two confinement spaces, we imagined "portals" (see Figure 17) that work as mechanisms of explicit interaction: the user can use them to circulate between two rooms.

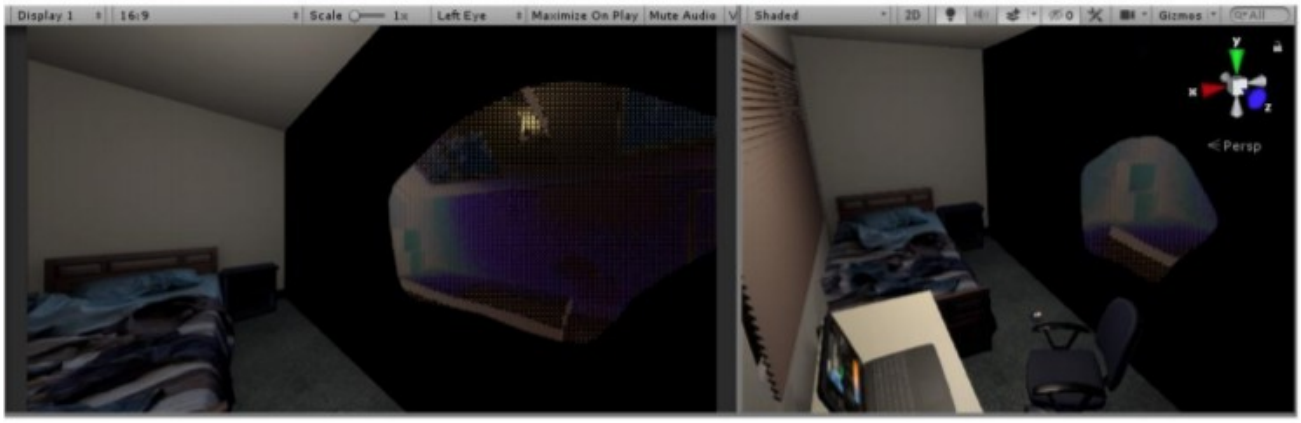


Figure 17. Creating virtual portals to travel between rooms.

Seeking to evoke the passage of time as a factor of realism in confined spaces, defined for each different color temperatures and sources of brightness in the room (see Figure 18), both indoors and outdoors. Dynamics of light intensity and source displacement were created to evoke everyday life.

Figure 18. Creation of light dynamics to accentuate the feeling of the passing of time.

All of the aspects described as part of the creation process are connected around the concept of experimental congruence, that is to say, they contribute in the creation of an integrated system of visual attributes whose common function is to create the sensation of a coherent virtual world that simplifies the user's immersions, forgetting the artificial features of the experience. Having explicit interactions with the environment is a possibility that is added to the objective, in order to facilitate the sensation of physical presence in the virtual space.

8.5.2. The design of implicit interactions and the immersion map

The lighting, chromatic, sonorous and musical aspects of the exterior scene are also governed by the experimental congruence previously described, but they also establish a range of attributes that can be the object of implicit interactions between the user and the environment. Conditions can be defined in which the corporal behaviour of the user in the interior of the virtual space could trigger changes in the environment (e.g changes in the waves, wind, etc.). By not being aware that their behaviour is the source of the changes, we can examine the way in which these environmental fluctuations modify their intellectual and emotional response. Thanks to the implicit interactions, we will be able to explore favorable conditions to trigger behaviors of intellectual and/or emotional immersion.

We defined an experimental prototype where one of the scenes is a boat trip. We chose this scenery to stage a global movement. The user could move from the previously described closed space to the ship deck, where they could walk, accessing the bow, port and starboard. The boat would advance through different landscapes, passing through different times of the day and weather conditions (see Figure 19).

The programming of these interactions supposes a previous classification of the user's behaviors (what s/he does), both at a temporal level (at which moment of the trip) and spatially (where on the deck s/he is). Thanks to the user's position we can define in what moments the interactions will occur. We will then be able to generate a sort of "immersion map": invent a visualization mode that combines those places and moments, within a temporary flow of the boat trip, which allows us to establish comparisons between users. This allows us to trace immersion trajectories for each user and establish comparisons between groups of users. If we validate this "immersion map" as a heuristic tool, we can carry out preliminary trials with users in different phases of the production process. This allows us, for example, to compare the before and after results of the incorporation of virtual characters, with the final purpose of understanding how the social interaction manner in virtual environments can generate relevant changes in immersive sensations and the presence of users.

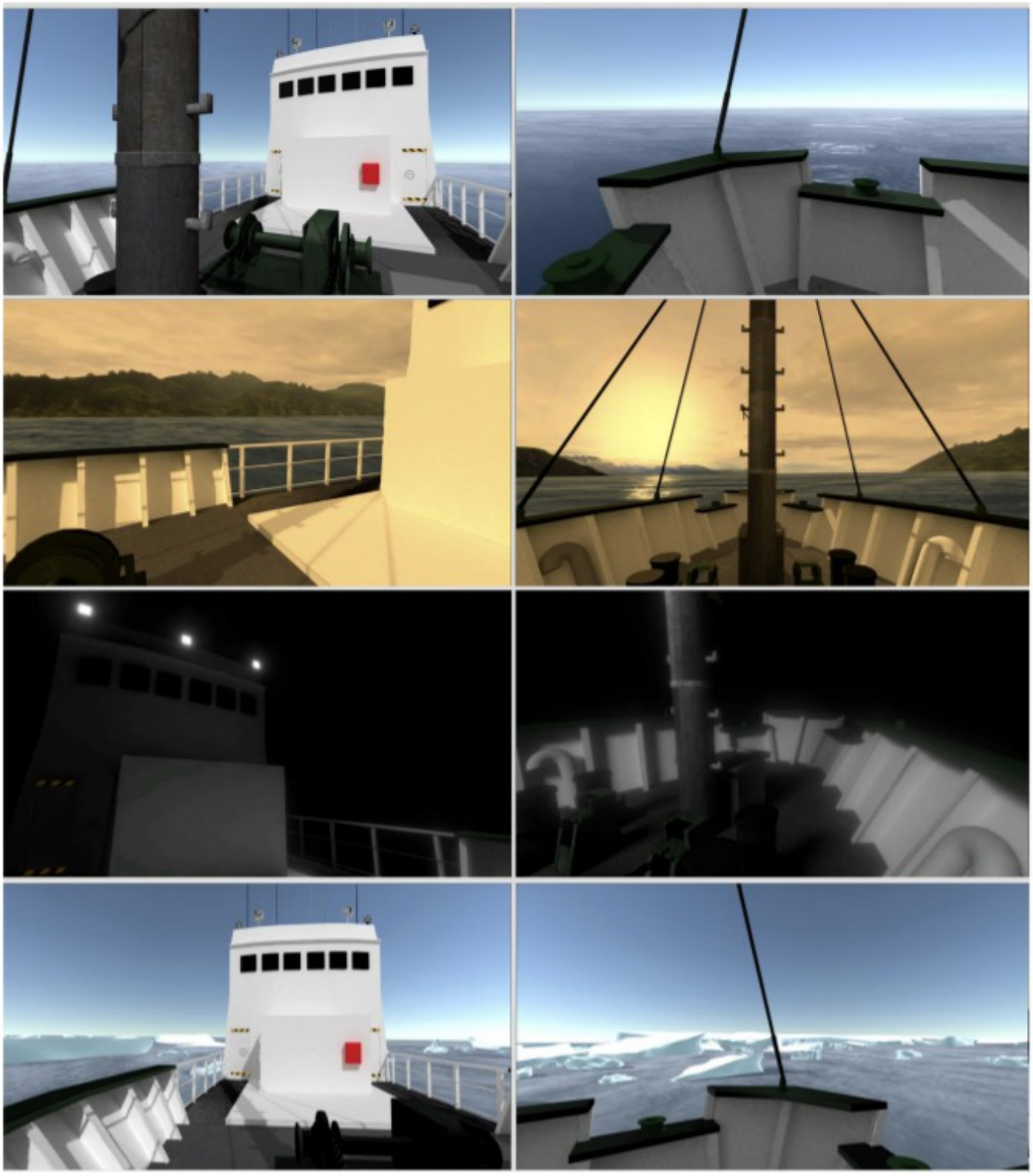


Figure 19. Boat trip during four different times of the day: daytime, sunset, night, sunrise.

8.6. Conclusive remarks

Immersion in virtual reality can be understood as a process of modulation between filmic immersion and interaction that requires a transdisciplinary approach to be able to reconcile the knowledge involved, so that audiovisual creation can be put at the service of research, and vice versa, thus favoring the creation of new knowledge.

The problems linked to the discomforts of immersive experiences can be approached from a transdisciplinary perspective that allows combining artistic, scientific and technological knowledge for various purposes: identifying factors causing discomfort associated with virtual movement, testing detection mechanisms for problematic audiovisual factors during the temporality of the contents and validate solutions for the improvement of the immersive experience.

In this chapter we present the research conclusions. In general terms, it is stated that NPR triggers specific ocular behaviors that can have a real effect on the way of perceiving a cinematic narrative. The transdisciplinary interest of this thesis lies in the fact that the concept of "film narrative" is not something exclusive to cinema, but that can also be conceived as a set of audiovisual procedures applicable to virtual reality and video games. The fundamental idea is that NPR post-production provides a visual style that can modify the way of perceiving a narrative, beyond the audiovisual format where it is applied.

The empirical research process is made up of three stages: (1) we questioned about the possible influence of NPR on cinematic perception in 2D cinema and stereoscopic 3D photography; then (2) we were interested in the ocular, attentional and pupillary behavior of the viewer faced to a film produced in stereoscopic 3D format with NPR post-production, and finally (3) we projected these results to study their application in the production of stereoscopic 3D virtual reality.

The first stage began with the general question about the influence of NPR on conventional film perception, that is, with 2D images projected on a large screen. We immediately apply this question to 3D photographs, as a way to prepare for the next stage: the study of NPR in 3D cinema. The main goal of this first stage is to deepen the compatibility between photographic stereoscopy and NPR: how our perception of 3D is modified with non-photorealistic processing.

In [FMCOS17], and in chapter 5, we argue that: NPR should not affect the narrative understanding of 2D cinema, since users are able to follow the story, through a main recognition of the character's identity (that is, his body movement and his voice). On the other hand, it was revealed that, in the transition from NPR in 2D to 3D cinema, the visual abstraction of NPR, fused with the optical hyperrealism of stereoscopy, does not turn out to be contradictory procedures, despite the scanning process of depth.

The use of NPR in cinema means a challenge to the global tradition of photorealism. The common sense of our visual culture tells us that NPR and stereoscopy are destined to be separated. In [FMCOS17] we suggest that this combination is not a perceptual contradiction. Our experimental results in 2D cinema indicate that narrative comprehension is not significantly affected by NPR. In 3D cinema we conducted a trial study that produced preliminary observations. These observations will be used in future research as the basis for a user study aimed at identifying NPR stylizations that may preserve the binocular depth of stereoscopy. Our current work indicates that a combination of the expressive quality of NPR with the immersive effect of stereoscopic cinema could produce a new form of augmented narrative.

In order to verify this first result, which points to the effect of perceptual modifications produced by NPR, we elaborated an experimental design based on the use of eye-tracking to collect data about ocular behavior of spectators. This allowed us to analyze ocular behavior in two dimensions: attentional and pupillary reactivity to NPR.

We concluded that a NPR filter associated with a high level of abstraction could trigger an adaptive ocular response that may be described as perceptual anxiety: a lower exploration of the scene and a greater focus on the more expressive elements of the scene: bodies and faces of the characters. This ocular behavior, triggered by NPR, could benefit cognitive immersion, especially in cinematographic genres (fantasy, horror, suspense, science fiction) that use visual abstraction and cognitive uncertainty as an aesthetic effect to favor greater involvement of viewers in their narrative movie.

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, in 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 less. 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, in which the 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.

The comparative analysis of ocular behavior under different visual styles opens up new perspectives for the understanding of the concept of immersion. Attention and emotional vision markers can be used to implement an empirical method, based on eye-tracking data, to study immersion in virtual reality. In order to deepen this line of research, it will be necessary to adapt the methods of data analysis, for example, by incorporating quantitative methods of ocular data analysis to identify objects of visual interest in the three-dimensional space of immersive experiences. If we can correlate the way we see with what interests us in a scene, we could describe our immersion experience with greater precision.

The discussion allowed us to reflect on how to apply our results on NPR perception to the creation of audiovisual experiences in interactive virtual reality. We set out to conceive a narrative experience based on the thesis of our research, regarding the perceptual change generated by NPR in the spectator's ocular behavior and its effects on the film narrative. This means that, depending on the degree of visual abstraction of each scene, we could assume that visual attention could focus on the characters, or on the contrary, assign attentional resources to the visual study of the environment. Our approach, focused on the temporal development of immersion, allows us to associate the variations in ocular behavior to the cinematic narrative. We believe that this way of interpreting the relationship between post-production and film narrative perception can become the basis for research, development and production of immersive audiovisual content.

We have adopted a transdisciplinary approach that combines the contributions of disciplines such as cinema, computing, sound engineering, social sciences, neurosciences and education, in order to link, in a single effort of audiovisual creation and research, reflections on physical behavior and cognition in interactive virtual environments. We combined procedures such as script writing, interactive design, the creation of characters and spaces, the development of interactive software, sound design and image post-production, to study the audiovisual production stages as an integrated process where all of the aesthetical attributes will display meanings to the experience of immersion in cinematic virtual reality environments.

The cinematic narrative is a set of language codes born from the art of cinema, however, with the arrival of new broadcast formats, digital platforms and the vertiginous development of video games, it is possible to find new forms of use for the cinematic narrative codes. Our thesis sustains that the visual style modifies the ocular behavior in a way that favors the functioning of the film narrative. We believe that some of the processes that contribute to film narrative are also found in virtual reality stories: the characters evolve on stage, speak, interact with each other, establish physical contact, etc. What we call “cinematic events” linked to characters can be found in a very similar way in virtual reality. However, the semi-spherical three-dimensional format (the 360-degree scene) of virtual reality generates fundamental changes in another dimension of the film narrative: what we call "cinematographic procedures", such as camera movement and film edition. Some elements of the film narrative remain but others are deeply transformed in virtual reality. This is even clearer in the case of interactive virtual reality, where the viewer can move inside the audiovisual space and manipulate objects, actions that would be unthinkable in cinema. Virtual reality shares some of the codes of cinema but incorporates behaviors that are common in video games, such as the need to consider user behavior in all of the phases of content design and production.

The contribution of our study of visual perception lies in the understanding of the user-viewer: we propose a way of interpreting the relationship between visual style and immersion that allows us to predict, under certain conditions, possible behaviors of the user. Our proposal is to use our results in virtual reality production and transform our findings into user behavior expectations that serve to guide the production process. In other words, we believe in the development of immersive content based on user behavior, observable thanks to the use of eye-tracking; this point of view will permit us to combine audiovisual creation and scientific research in a way that contributes to the creation of new knowledge. Scientific insights will guide the design of the scenes and audiovisual knowledge, expressed in specific scenes, will allow to enlarge the scope of scientific hypotheses on user behavior.

This synergy between art and science is the basis of our transdisciplinary approach to the study of user immersion in interactive virtual environments.

The use of eye-tracking to study the impact of NPR in 3D cinema has provided us with evidence of great interest to explore and describe the sensation of cinematic immersion in virtual reality.

Even though the tests showed that our chosen NPR methods were compatible to stereoscopy, we also plan to apply the method of Du et al. [DMHG13] to evaluate the stereoscopic viewing comfort of the stylized images quantitatively. To broaden the study of NPR effects on film narrative, it will be necessary to incorporate and test new modalities of NPR processing, in such a way that we can compare different methods of NPR processing and study their respective perceptual effects, including the emotional dimension through the study of the pupil response. Finally, it will also be necessary to consider the importance of other variables that condition the scope of this study, such as the narrative style and the cinematographic genre of the films, the type of audiovi98sual visualization format (cinema, television, VR), the age of life and the context of film culture of the viewers considered as subjects for research.

However, as virtual reality is a different medium, with its own language codes in full emergency, it seems necessary to develop and expand our research approach. To be able to describe the particular way in which the film narrative is transformed in virtual reality, we will consider four main factors: a visual format (in 360) different from the photographic framing, an auditory experience (binaural or ambisonic) that modifies auditory perception depending on the position and orientation of the body, the ability to design interactions (explicit or implicit) between the user and their environment, and finally, the possibility that the audiovisual experience can generate physical discomfort that directly affect the quality of the immersion experience. The film narrative, when transformed by these factors, becomes a "virtual narrative".

The immersion is no longer solely mental, it is now completely bodily. It is for this reason that it is necessary to design a new research epistemology, capable of combining different technical, artistic, scientific and technological knowledge, in a new approach where audiovisual production and the study of user behavior become inseparable. This is the main perspective for future work derived from this doctoral research: we are conceiving a research approach that allows us to use eye-tracking in virtual reality to be able to understand, once

again, the relationship between visual style and ocular behavior. But since it is not about film narratives in their traditional sense, we are no longer talking about cinema but about virtual reality. We must study the emerging narrative codes —the virtual narrative— to build a new interpretive framework that allows us to interpret the data of ocular behavior.

To guide future research that arises from this thesis, we have created the Virtual Realities Lab (VR-LAB) at the University of Chile: an academic space for international collaboration between artists and scientists aimed at developing a transdisciplinary methodology to incorporate eye-tracking and non-photorealistic rendering on virtual reality research.

The experience of having lived several months of confinement due to the Covid-19 pandemic allowed me to observe a significant development of virtual environments called VR-Rooms or Virtual Worlds. Many film festivals have chosen to use virtual environments to promote social encounters and the exhibition of content. Several video games have entered the virtual reality format, rethinking their narratives for this new platform. The number of immersive films in virtual reality continues to increase, as well as the development of virtual reality applications in social, educational, health or industrial contexts. Although this pandemic can be fought through vaccines and new forms of health worldwide, it seems very likely that the expansion of these new virtual platforms will not stop in the short term. So there is a plausible scope for our future work around NPR, eye-tracking and virtual reality. It is for this reason that the main research field in our VR-LAB is the combination of arts and sciences to apply this knowledge about ocular behavior to the production of narrative video games, cinematic immersive experiences and interactive virtual reality simulations.

The next step in our research is to use our interactive virtual reality prototype and design an experiment where we can evaluate, with the use of eye-tracking, the ocular behavior inside an immersive space. This is the future work of our transdisciplinary research.

REFERENCES

[AA10] Andre Almeida and Heitor Alvelos. An interactive documentary manifesto. In Ruth Aylett, Mei Lim, Sandy Louchart, Paolo Petta, and Mark Riedl, editors, Interactive Storytelling, pages 123-128. Springer, 2010.

[AF17] Zahid Akhtar and Tiago H. Falk. Audio-Visual Multimedia Quality Assessment: A Comprehensive Survey. IEEE Access, 5, 21090–21117, 2017.

[Agr09] Amit Agrawal. Non-photorealistic Rendering: Unleashing the Artist's Imagination [Graphically Speaking]. IEEE Computer Graphics and Applications, 29(4): 81–85, 2009.

[Ans] ANSES. Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail. https://www.anses.fr

[Ans14] ANSES. 3D technologies and eyesight: use not recommended for children under the age of six, use in moderation for those under the age of 13. https://goo.gl/jOMm5o, 2014.

[APA18] American Physiological Society. Motion sickness vs. cybersickness: Two different problems or the same condition? findings of a new study contradict previous research. ScienceDaily. http://www.sciencedaily.com/releases/2018/10/181023085654.htm, 2018.

[Aum92] Jacques Aumont. Du visage au cinéma. Editions de l'Etoile, 1992.

[Bar04] Juddy Barrett. Side effects of virtual environments: A review of the literature. Defense Science and Technology Organisation, 2004.

[BBI13] Stacy Balk, Dakota Bertola, and Vaughan Inam. Simulator sickness questionnaire: Twenty years later. In Proceedings of the Seventh International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, pages 257–263, 2013.

[BBRQGA08] Rosa M. Baños, Cristina Botella, Isabel Rubió, Soledad Quero, Azucena García-Palacios, and Mariano Alcañiz. Presence and Emotions in Virtual Environments: The Influence of Stereoscopy. CyberPsychology & Behavior, 11(1): 1–8, 2008.

[BC04] Emily Brown and Paul Cairns. A grounded investigation of game immersion. In CHI '04 Extended Abstracts on Human Factors in Computing Systems, pages 1297–1930, 2004.

[BC04] Emily Brown and Paul Cairns. A grounded investigation of game immersion. In CHI'04: Extended Abstracts on Human Factors in Computing Systems, pages 1297–1300. ACM, 2004.

[BJ19] Serim Baris and Giulio Jacucci. Explicating "Implicit Interaction": An Examination of the Concept and Challenges for Research. In CHI'19: Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems, pages 1–16. ACM Press, 2019.

[BLCC08] Alexandre Benoit, Patrick Le Callet, Patrizio Campisi, and Romain Cousseau. Quality Assessment of Stereoscopic Images. EURASIP Journal on Image and Video Processing, 1–13, 2008.

[BM09] Simon Barthelmé and Pascal Mamassian. Evaluation of objective uncertainty in the visual system. Evaluation of Objective Uncertainty in the Visual System. PLoS Comput Biol, 5(9):

e1000504, 2009.

[BMEL08] Margaret M. Bradley, Laura Miccoli, Miguel A. Escrig, and Peter J. Lang. The pupil as a measure of emotional arousal and autonomic activation. Psychophysiology, 45(4): 602–607, 2008.

[BRC19] Tobias Benz, Bernhard Riedl, and Lewis Chuang. Projection Displays Induce Less Simulator Sickness than Head-Mounted Displays in a Real Vehicle Driving Simulator. In Christian JANSSEN and Stella DONKER, editors, AutomotiveUI '19: 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, pages 379-387. Association for Computing Machinery, 2019.

[BRR07] Stephane Bouchard, Geneviève Robillard, and Patrice Renaud. Revising the factor structure of the Simulator Sickness Questionnaire. Annual Review of CyberTherapy and Telemedicine, 2007.

[BS16] Giovanni Bertolini and Domink Straumann. Moving in a Moving World: A Review on Vestibular Motion Sickness. Frontiers in Neurology, 7(14), 1–11, 2016.

[CD14] Giorgio Conte and Massimo De Marchi. Non-Photorealistic Rendering: From a general view to suggestive contours, 2014.

[CDFKMS06] Forrester Cole, Doug DeCarlo, Adam Finkelstein, Kenrick Kin, Keith Morley, and Anthony Santella. Directing gaze in 3D models with stylized focus. In Eurographics Symposium on Rendering, pages 377–387, 2006.

[CDLK19] Cassidy Curtis, Kevin Dart, Theresa Latzko, and John Kahrs. Non-photorealistic animation for immersive storytelling. In C. Kaplan, A. Forbes, and S. DiVerdi, editors, The 8th ACM/EG Expressive Symposium, pages 1–10. The Eurographics Association, 2019.

[CKY20] Eunhee Chang, Hyun Taek Kim, and Byounghyun Yoo. Virtual Reality Sickness: A Review of Causes and Measurements. International Journal of Human–Computer Interaction, 1–25, 2020.

[CN08] Manuel G. Calvo and Lauri Nummenmaa. Detection of emotional faces: salient physical features guide effective visual search. Journal of Experimental Psychology: General, 137(3): 471–494, 2008.

[Cro08] Karen Cronin. Transdisciplinary research (TDR) and sustainability, Overview report prepared for the Ministry of Research, Science and Technology (MoRST). Environmental Science and Research (ESR) Ltd, 2008.

[DB10] Frédéric Devernay and Paul Beardsley. Stereoscopic cinema. In Geometry Processing for 3-D Cinematography, pages 11–51. Springer, 2010.

[DD18] Mark Dennison and Michael D'Zmura. Effects of unexpected visual motion on postural sway and motion sickness. Applied Ergonomics, 71: 9–16, 2018.

[Ded09] Chris Dede. Immersive interfaces for engagement and learning. Science, 323(5910): 66–69, 2009.

[Del12] Rafael Del Villar. Navegación por Internet: Percepción y cognición en población adulta del Gran Santiago. Proyecto FONDECYT N° 1120064, 2012.

[Del15] Rafael Del Villar. Veridicción en condiciones de simplicidad/ complejidad cognitiva. In Marcello Serra and Oscar Gómez, editors, Transparencia y secreto, pages 87–97. Visor Libros, 2015.

[DMHG13] Song-Pei Du, Belen Masia, Shi-Min Hu, and Diego Gutierrez. A metric of visual comfort for stereoscopic motion. ACM Transactions on Graphics, 32(6): 222:1–222:9, 2013.

[DSS18] Natalia Dużmańska, Paweł Strojny, and Agnieszka Strojny. Can Simulator Sickness be Avoided? A review on Temporal Aspects of Simulator Sickness. Frontiers in Psychology, 9, 2018.

[FGMDYMMS18] Victor Fajnzylber, Larry González, Pedro Maldonado, Rafael Del Villar, Rodrigo Yáñez, Samuel Madariaga, Milán Magdics, and Mateu Sbert. Augmented film narrative by use of non-photorealistic rendering. In 2017 International Conference on 3D Immersion (IC3D), pages 1–8. IEEE, 2018.

[FMCOS17] Victor Fajnzylber, Milán Magdics, Macarena Castillo, Constanza Ortega, and Mateu Sbert. From 2D to 3D: A Case Study of NPR and Stereoscopic Cinema. In International Symposium on Smart Graphics, pages 87–98. Springer, 2017.

[FMDMVSM19] Victor Fajnzylber, Samuel Madariaga, Rafael Del Villar, Pedro Maldonado, Diego Vargas, Aria Serra, Milan Magdics, and Mateu Sbert. Pupillary Reactivity to Non-Photorealistic Rendering: A Case Study of Immersion in 3D Cinema. In 2018 International Conference on 3D Immersion (IC3D), pages 1–8. IEEE, 2019.

[Fra] FrameForge. Award-Winning Storyboarding and Previs Software. https://www.frameforge.com

[Fre00] Walter Freeman. Neurodynamics: An exploration in mesoscopic brain dynamics. Springer Science & Business Media, 2000.

[FWMS19] Victor Fajnzylber, Andrea Wenner, Javier Moyano, and Mateu Sbert. From NPR to VR: tracking ocular behavior in immersive virtual reality. Communication Papers, 8(17): 21–35, 2019.

[Gen10] Weidong Geng. Introduction. In The Algorithms and Principles of Non-Photorealistic Graphics: Artistic Rendering and Cartoon Animation. Springer Science & Business Media, 2010.

[GN13] Dan W. Grupe and Jack B. Nitschke. Uncertainty and anticipation in anxiety: an integrated neurobiological and psychological perspective. Nature Reviews Neuroscience, 14(7): 488–501, 2013.

[GRG04] Bruce Gooch, Erik Reinhard, and Amy Gooch. Human Facial Illustrations: Creation and Psychophysical Evaluation. ACM Transactions on Graphics (TOG), 23(1): 27–44, 2004.

[GW02] Amy Ashurst Gooch and Peter Willemsen. Evaluating space perception in NPR immersive environments. In NPAR '02: Proceedings of the 2nd International Symposium on Non-photorealistic Animation and Rendering, pages 105–110. ACM, 2002.

[HBL18] Robert R. Henderson, Margaret M. Bradley, Peter J. Lang. Emotional Imagery and Pupil Diameter. Psychophysiology, 5(6): 1–7, 2018.

[Her10] Aaron Hertzmann. Non-photorealistic Rendering and the Science of Art. In NPAR ’10: Proceedings of the 8th International Symposium on Non-Photorealistic Animation and Rendering, pages 147–157, 2010.

[HH63] Richard Held and Alan Hein. Movement-produced stimulation in the development of visually guided behavior. Journal of Comparative and Physiological Psychology, 56(5), 872–876, 1963.

[HHD04] Michael Haller, Christian Hanl, and Jeremiah Diephuis. Non-photorealistic rendering techniques for motion in computer games. ACM Computers in Entertainment (CIE), 2(4): 1–10, 2004.

[HI16] Rostam Affendi Hamzah and Haidi Ibrahim. Literature survey on stereo vision disparity map algorithms. Journal of Sensors, 8742920:1–8742920:23, 2016.

[HMHLS03] Nick Halper, Mara Mellin, Christoph S. Herrmann, Volker Linneweber, and Thomas Strothotte. Psychology and Non-Photorealistic Rendering: The Beginning of a Beautiful Relationship, In Mensch & Computer 2003, pages 277–286, 2003.

[Hot51] Harold 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.

[HVC12] F. Honbolygó, L. Veller, and V. Csépe. Ventriloquism aftereffect in a virtual audio-visual environment. In 2012 IEEE 3rd International Conference on Cognitive Infocommunications (CogInfoCom), pages 475–478. IEE, 2012.

[JL08] Wendy Ju and Larry Leifer. The design of implicit interactions: Making interactive systems less obnoxious. Design Issues, 24(3): 72–84, 2008.

[Ju15] Wendy Ju. The Design of Implicit Interactions. Morgan & Claypool Publishers, 2015.

[KA19] Mara Kaufeld and Thomas Alexander. The impact of motion on individual simulator sickness in a moving Base VR simulator with head-mounted display (HMD). In International Conference on Human-Computer Interaction, pages 461–472, Springer, 2019.

[KCWI13] Jan Eric Kyprianidis, John Collomosse, Tinghuai Wang, and Tobias Isenberg. State of the ‘art’: A taxonomy of artistic stylization techniques for images and video. IEEE Transactions on Visualization and Computer Graphics, 19(5): 866–885, 2013.

[KD08] Jan Eric Kyprianidis and Jürgen Döllner. Image abstraction by structure adaptive filtering. In EG UK Theory and Practice of Computer Graphics, pages 51–58, 2008.

[KDK10] Robert S. Kennedy, Julie Drexler, Robert C. Kennedy. Research in visually induced motion sickness. Applied Ergonomics, 41(4): 494–503, 2010.

[KHP20] Jiwon Kim, Jihong Hwang, and Taezoon Park. Effect of Motion Cues on Simulator Sickness in a Flight Simulator. In International Conference on Human-Computer Interaction, pages 493–506, Springer, 2020.

[KJK17] Mingyu Kim, Changyu Jeon, and Jinmo Kim. A study on immersion and presence of a portable hand haptic system for immersive virtual reality. Sensors, 17(5), 1141, 1–18, 2017.

[KKO15] Mikko Kytö, Kenta Kusumoto, and Pirkko Oittinen. The ventriloquist effect in augmented reality. In IEEE International Symposium on Mixed and Augmented Reality (ISMAR), pages 49–53. IEEE, 2015.

[KLKL13] Yongjin Kim, Yunjin Lee, Henry Kang, and Seungyong Lee. Stereoscopic 3d line drawing. ACM Transactions on Graphics, 32(4): 57:1–57:13, 2013.

[KM06] Atsushi Kasao and Kazunori Miyata. Algorithmic Painter: A NPR method to generate various styles of painting. The Visual Computer, 22(1): 14–27, 2006.

[Kol95] Eugenia Kolasinski. Simulator Sickness in Virtual Environments. U.S. Army Research Institute for the Behavioral and Social Sciences, 1995.

[KPNM19] M. P. Pavan Kumar, B. Poornima, H. S. Nagendraswamy & C. Manjunath. A comprehensive survey on non-photorealistic rendering and benchmark developments for image abstraction and stylization. Iran Journal of Computer Science, 2: 131–165, 2019.

[KRHC15] Behrang Keshavarz, Bernhard E. Riecke, Lawrence J. Hettinger, and Jennifer L. Campos. Vection and visually induced motion sickness: how are they related?. Frontiers in Psychology, 6:472.

[LaV20] Steven LaValle. Virtual Reality. Cambridge University Press, 2020.

[LB14] Dajin Li and Chengjie Bai. A Simple Artistic Rendering Method for Stereoscopic Images. In J. Keyser, Y. J. Kim, and P. Wonka, editors, Pacific Graphics, 2014.

[LFHI09] Marc Lambooij, Marten Fortuin, Ingrid Heynderickx and Wijnand IJsselsteijn. Visual discomfort and visual fatigue of stereoscopic displays: a review. Journal of Imaging Science and Technology, 53(3): 030201–030201-14, 2009.

[LHCLLC11] Chun-Wei Liu, Tz-Huan Huang, Ming-Hsu Chang, Ken-Yi Lee, Chia-Kai Liang, and Yung-Yu Chuang. 3D cinematography principles and their applications to stereoscopic media processing. In MM '11: Proceedings of the 19th ACM international conference on Multimedia, pages 253–262. ACM, 2011.

[Lin91] Jianhua Lin. Divergence measures based on the Shannon entropy. IEEE Transactions on Information Theory, 37(1): 145–151, 1991.

[LL20] Le Quang Van, Le Quan Van, Nishimaru Hiroshi, Matsumoto Jumpei, Takamura Yusaku, Hori Etsuro, Maior Rafael S., Tomaz Carlos, Ono Taketoshi, Nishijo Hisao. A Prototypical Template for Rapid Face Detection Is Embedded in the Monkey Superior Colliculus. In

Frontiers in Systems Neuroscience, 14, 2020.

[LLMS15] Lester C. Loschky, Adam M. Larson, Joseph P. Magliano, and Tim J. Smith. What would jaws do? the tyranny of film and the relationship between gaze and higher-level narrative film comprehension. PLOS ONE, 10(11): e0142474, 2015.

[LMD13] Hua Li, David Mould, and Jim Davies. Structure and aesthetics in non-photorealistic images. In Graphics Interface, pages 181–188, 2013.

[LRBL12] Pierre Lebreton, Alexander Raake, Marcus Barkowsky, Patrick Le Callet. Evaluating Depth Perception of 3D Stereoscopic Videos. IEEE Journal of Selected Topics in Signal Processing, 6(6): 710–720, 2012.

[McC94] Scott McCloud. Understanding Comics: The Invisible Art. Harper Collins & Kitchen Sink Press, 1994.

[McM03] Alison McMahan. Immersion, engagement, and presence: A method for analyzing 3-D video games. In The Video Game Theory Reader, pages 67–87. Routledge, 2003.

[Mes05] Daniel Mestre. Immersion and Presence. Movement & Perception, CNRS & University of the Mediterranean, 2005.

[MGVP04] Arpi Minassian, Eric Granholm, Steven Verney, and William Perry. Pupillary dilation to simple vs. complex tasks and its relationship to thought disturbance in schizophrenia patients. International Journal of Psychophysiology, 52(1): 53–62, 2004.

[MHG14] John P. McIntire, Paul R. Havig, and Eric E. Geiselman. Stereoscopic 3D displays and human performance: A comprehensive review. Displays, 35(1): 18–26, 2014.

[MML11] Regan L. Mandryk, David R. Mould, Hua Li. Evaluation of emotional response to non-photorealistic images. In NPAR ’11: Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Non-Photorealistic Animation and Rendering, 2011.

[MML12] David Mould, Regan L. Mandryk, and Hua Li. Emotional response and visual attention to non-photorealistic images. Computers & Graphics, 36(6): 658–672, 2012.

[MSGS13] Milán Magdics, Catherine Sauvaget, Rubén J. Garcia, and Mateu Sbert. Post-processing NPR effects for video games. In VRCAI '13: Proceedings of the 12th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry, pages 147–156, 2013.

[MWHN18] Alireza Mazloumi, Frederick Walker, Deborah Hodgson, and Eugene Nalivaiko. A comparative study of cybersickness during exposure to virtual reality and “classic” motion sickness: are they different? Journal of Applied Physiology, 125(6), 1670–1680, 2018.

[NAK13] Lesley Northam, Paul Asente, and Craig S. Kaplan. Stereoscopic 3D image stylization. Computers & graphics, 37(5): 389–402, 2013.

[NGJT13] Kristian Nymoen, Rolf Inge Godøy, Alexander Refsum Jensenius, and Jim Torresen. Analyzing correspondence between sound objects and body motion. ACM Transactions on Applied Perception, 10(2): 9:1–9:22, 2013.

[ODBK18] Mathias Oberhauser, Daniel Dreyer, Reinhard Braunstingl, and Ioana Koglbauer. What's real about virtual reality flight simulation? Comparing the fidelity of a virtual reality with a conventional flight simulation environment. Aviation Psychology and Applied Human Factors, 8(1), 22–34, 2018.

[PCSRVC00] Emilee Patrick, Dennis Cosgrove, Aleksandra Slavkovic, Jennifer A. Rode, Thom Verratti, and Greg Chiselko. Using a large projection screen as an alternative to head-mounted displays for virtual environments. In CHI '00: Proceedings of the SIGCHI conference on Human Factors in Computing Systems, pages 478-485. ACM, 2000.

[PCTRS20] Thiago Porcino, Esteban Clua, Daniela Trevisan, Érick Rodrigues, and Alexandre Silva. Automatic Recommendation of Strategies for Minimizing Discomfort in Virtual Environments. In IEEE 8th International Conference on Serious Games and Applications for Health. IEEE, 2020.

[PR16] Akshay Gadi Patil and Shanmuganathan Raman. Automatic Content-aware Non-Photorealistic Rendering of Images. arXiv:1604.01962v4, 1–6, 2016.

[Rad02] Pablo M. Rademacher. Measuring the perceived visual realism of images (Doctoral dissertation). University of North Carolina at Chapel Hill, 2002.

[RC15] Kyung Ryu and Sook-Hee Cho. The effects of visual information on anxiety and uncertainty in elderly patients after the total knee arthroplasty. Journal of Muscle and Joint Health, 22(1): 48–56, 2015.

[RD09] Niall Redmond and John Dingliana. Influencing user attention using real-time stylised rendering. In EG UK Theory and Practice of Computer Graphics, pages 173–180, 2009.

[RKD10] Christian Richardt, Jan Eric Kyprianidis, and Neil Dodgson. Stereo coherence in watercolour rendering. In Symposium on Non-Photorealistic Rendering and Animation 2010, 2010.

[RO16] Lisa Rebenitsch and Charles Owen. Review on Cybersickness in Applications and Visual Displays. Virtual Reality, 20(2), 101–125, 2016.

[Sch00] Albrecht Schmidt. Implicit human computer interaction through context. Personal and Ubiquitous Computing, 4(2), 191–199, 2000.

[SCH09] Puneet Sharma, Faouzi Alaya Cheikh, and Jon Yngve Hardeberg. Face saliency in various human visual saliency models. In Proceedings of 6th International Symposium on Image and Signal Processing and Analysis, pages 327–333, 2009.

[SD04] Anthony Santella and Doug DeCarlo. Visual interest and NPR: An evaluation and manifesto. In NPAR '04: Proceedings of the 3rd International Symposium on Non-photorealistic Animation and Rendering, pages 71–150, ACM, 2004.

[SG04] E. Stavrakis and M. Gelautz. Image-based stereoscopic painterly rendering. In EGSR'04: Proceedings of the Fifteenth Eurographics conference on Rendering Techniques, pages 53–60, 2004.

[SGCDPSGBG20] Jean-Christophe Servotte, Manon Goosse, Suzanne Campbell, Nadia Dardenne, Bruno Pilote, Ivan Simoneau, Michèle Guillaume, Isabelle Bragard, and Alexandre Ghuysen. Virtual Reality Experience: Immersion, Sense of Presence, and Cybersickness. Clinical Simulation in Nursing, 38, 35–43, 2020.

[SH06] Rezwan Sayeed and Toby Howard. State of the Art Non-Photorealistic Rendering (NPR) Techniques. In Louise M. Lever and Mary McDerby, editors, Theory and Practice of Computer Graphics 2006. The Eurographics Association, 2006.

[SH09] Geng Sun and Nick Holliman. Evaluating methods for controlling depth perception in stereoscopic cinematography. In A. J. Woods, N. S. Holliman, J. O. Merritt, editors, Stereoscopic Displays and Applications XX, 7237: 72370I. International Society for Optics and Photonics, 2009.

[SI20] Volkan Sevinca and Berkman Mehmet Ilker. Psychometric evaluation of Simulator Sickness Questionnaire and its variants as a measure of cybersickness in consumer virtual environments. Applied Ergonomics, 82, 2020.

[Sla03] Mel Slater. A note on presence terminology. Presence-Connect, 3(3): 1–5, 2003.

[SM13] Tim J. Smith and Parag Mital. Attentional synchrony and the influence of viewing task on gaze behavior in static and dynamic scenes. Journal of Vision, 13(8): 1–24, 2013. [SM15] S. K. G. e. Silva and F. d. L. d. S. N. Marques. Depth Perception Evaluation with Different Stereoscopic Techniques: A Case Study. In 2015 XVII Symposium on Virtual and Augmented Reality, pages 52–60, Sao Paulo, 2015.

[SM93] Barry E. Stein and M. Alex Meredith. The Merging of the Senses. MIT Press, 1993.

[Smi12] Tim J. Smith. The attentional theory of cinematic continuity. Projections, 6(1): 1–27, 2012.

[SN12] Junichiro Seyama and Ruth S Nagayama. Detection of a change in photorealism. Japanese Psychological Research, 54(4): 378–387, 2012.

[SOBENG16] Robert J. Snowden, Katherine R. O'Farrell, Daniel Burley, Jonathan T. Erichsen, Naomi V. Newton, and Nicola S. Gray. The pupil's response to affective pictures: Role of image duration, habituation, and viewing mode. Psychophysiology, 53(8): 1217–1223, 2016.

[Sol13] Angelo G. Solimini. Are there side effects to watching 3D movies? A prospective crossover observational study on visually induced motion sickness. PloS one, 8(2): e56160, 2013.

[SSv16] Mel Slater and Maria Sanchez-Vives. Enhancing our lives with immersive virtual reality. Frontiers in Robotics and AI, 3(74), 1–47, 2016.

[Sta08] Efstathios Stavrakis. Stereoscopic non-photorealistic rendering (PhD thesis). Vienna University of Technology, 2008.

[Sto17] William Stone. Psychometric Evaluation of the Simulator Sickness Questionnaire as a

Measure of Cybersickness. Graduate Theses and Dissertations, 2017.

[SvS05] Maria Sanchez-Vives and SLATER Mel. From presence to consciousness through virtual reality. Nature Reviews Neuroscience, 6(4), 332–339, 2005.

[SW97] Mel Slater and Sylvia Wilbur. A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments. Presence: Teleoperators and Virtual Environments, 6(6): 603–616, 1997.

[TCNN10] Daniel M. Tokunaga, Cléber G. Corrêa, Ricardo Nakamura, Fátima L. S. Nunes, and Romero Tori. Non-photorealistic rendering in stereoscopic 3D visualization. ACM SIGGRAPH 2010 Posters on - SIGGRAPH '10, 2010.

[TDMS14] Krzysztof Templin, Piotr Didyk, Karol Myszkowski, and Hans- Peter Seidel. Perceptually-motivated stereoscopic film grain. Computer Graphics Forum, 33(7): 349–358, 2014.

[TM98] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In Proceedings of the Sixth International Conference on Computer Vision, pages 839–846, 1998.

[TMF19] Hiroki Takada, Masaru Miyao, Sina Fateh. Stereopsis and Hygiene. Current Topics in Environmental Health and Preventive Medicine, 2019.

[Toc16] Yvonne Toczek. The Influence of Visual Realism on the Sense of Presence in Virtual Environments. Eindhoven, 2016.

[TOLLC17] Mary Anne Thompson, Susan Owen, Jan Lindsay, Graham Leonard, and Shane Cronin. Scientist and stakeholder perspectives of transdisciplinary research: Early attitudes, expectations, and tensions. Environmental Science & Policy, 74, 30–39, 2017.

[UDWB16] William Uricchio, Sue Ding, Sarah Wolozin, and Beyza Boyacioglu. Virtually There: Documentary Meets Virtual Reality. MIT Open Documentary Lab, the John D and Catherine T MacArthur Foundation and the Phi Centre, 2016.

[UH08] Kazuhiko Ukaia and Peter A. Howarth. Visual fatigue caused by viewing stereoscopic motion mages: Background, theories, and observations. Displays, 29(2): 106–116, 2008.

[Uni] Unity Technologies. Unity real-time 3D development platform. https://unity.com/

[USS11] Tamás Umenhoffer, László Szécsi, and László Szirmay- Kalos. Hatching for motion picture p roduction. In Computer Graphics Forum, 30(2), pages 533–542. Wiley Online Library, 2011.

[UW11] Hiroyasu Ujike and Hiroshi Watanabe. Effects of stereoscopic presentation on visually induced motion sickness. Stereoscopic Displays and Applications XXII, 2011.

[Var95] Francisco Varela. Conocer: Las ciencias cognitivas: tendencias y perspectivas. Cartografía de las ideas actuales. Gedisa, 1995.

[Var99] Francisco Varela (1999). The specious present: A neurophenomenology of time consciousness. In Jean Petitot, Francisco Varela, Bernard Pachoud, and Jean-Michael Roy, editors, Naturalizing Phenomenology: Issues in Contemporary Phenomenology and Cognitive Science, pages 266–314. Stanford University Press, 1999.

[VCI19] Ivan Viola, Min Chen, and Tobias Isenberg. Visual Abstraction. arXiv-1910.03310, 2019.

[VdG00] Jean Vroomen and Beatrice de Gelder. Sound enhances visual perception: cross-modal effects of auditory organization on vision. Journal of Experimental Psychology: Human Perception and Performance, 26(5): 1583–1590, 2000.

[Vid1] Sample Video 1. Comparison of NPR and PR styles used in the experiment. https://vimeo.com/234386284, 2017.

[Vid2] Sample Video 2. Comparison of eye tracking results. https: //vimeo.com/234435709, 2017.

[VR04] Francisco J. Varela and Franz Regnot. Quel savoir pour l'éthique?: Action, sagesse et cognition. La Découverte/Poche, 2004.

[VTM10] Valentijn T. Visch, Ed S. Tan, and Dylan Molenaar. The emotional and cognitive effect of immersion in film viewing. Cognition & Emotion, 24(8): 1439–1445, 2010.

[Web15] Michael A. Webster. Visual adaptation. Annual Review of Vision Science, 1:547–567, 2015.

[Wid19] Joseph Wideström. The Transdisciplinary Nature of Virtual Space. In International Conference on Augmented Reality, Virtual Reality and Computer Graphics, p. pages 186–202. Springer, 2019.

[Win06] Holger Winnemöller. Perceptually-motivated non-photorealistic graphics (Doctoral dissertation). Northwestern University, 2006.

[WKO12] Holger Winnemöller, Jan Eric Kyprianidis, Sven C. Olsen. XDoG: An eXtended difference-of-Gaussians compendium including advanced image stylization. Computers & Graphics, 36(6): 720–753, 2012.

[WOG06] Holger Winnemöller, Sven C. Olsen, and Bruce Gooch. Real-time video abstraction. In SIGGRAPH '06: ACM SIGGRAPH 2006 Papers, pages 1221–1226, 2006.

[WS08] Louise Whiteley and Maneesh Sahani. Implicit knowledge of visual uncertainty guides decisions with asymmetric outcomes. Journal of Vision, 8(3): 1–15, 2008.

[WS15] Christopher J. Wilson and Alessandro Soranzo. The Use of Virtual Reality in Psychology: A Case Study in Visual Perception. Computational and Mathematical Methods in Medicine, 1–7, 2015.

[WS98] Bob G. Witmer and Michael J. Singer. Measuring presence in virtual environments: A presence questionnaire. Presence, 7(3): 225–240, 1998.

[ZMM19] Katja Zibrek, Sean Martin, and Rachel McDonnell. Is Photorealism Important for Perception of Expressive Virtual Humans in Virtual Reality?. ACM Transactions on Applied Perception (TAP), 16(3): 1–19, 2019.