# **RESEARCH CENTRE**

Inria Centre at Université Côte d'Azur

# 2024 ACTIVITY REPORT

Project-Team BIOVISION

Biologically plausible Integrative mOdels of the Visual system : towards synergIstic Solutions for visually-Impaired people and artificial visiON

DOMAIN Digital Health, Biology and Earth

THEME

Computational Neuroscience and Medicine



# Contents

Project-Team BIOVISION 1					
1	1 Team members, visitors, external collaborator	2			
2	2 Overall objectives	Overall objectives 3			
3	<ul> <li><b>3 Research program</b></li> <li>3.1 Axis 1 - Modeling the retina and the primar</li> <li>3.2 Axis 2 - Diagnosis, rehabilitation, and low-w</li> <li>3.3 Axis 3 - Visual media analysis and creation.</li> </ul>	4         4           y visual system.         4           ision aids.         4           .         7			
4	<ul> <li>4 Application domains</li> <li>4.1 Applications of low-vision studies</li> <li>4.2 Applications of vision modeling studies</li> <li>4.3 Applications of visual media analysis and complexity</li> </ul>	7 			
5	5 Social and environmental responsibility	Social and environmental responsibility 8			
6	<ul><li>6 Highlights of the year</li><li>6.1 Awards</li></ul>	<b>9</b> 			
7	<ul> <li>7 New software, platforms, open data</li> <li>7.1 New software</li> <li>7.1.1 Macular</li> <li>7.1.2 GUsT-3D</li> <li>7.1.3 PTVR</li> <li>7.1.4 Tatoovi</li> <li>7.2 New platforms</li> <li>7.2.1 CREATTIVE3D platform for navigati</li> <li>7.3 Open data</li> </ul>	9 9 9 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0			
8	<ul> <li>8 New results</li> <li>8.1 Modeling the retina and the primary visual</li> <li>8.1.1 Temporal pattern recognition in reinhibitory synapses</li> <li>8.1.2 How does the inner retinal networ computational study</li> <li>8.1.3 A chimera model for motion anticipation 8.1.4 Optical Flow Estimation pre-training</li> </ul>	13system13sinal ganglion cells is mediated by dynamical			
	<ul> <li>8.2 Diagnosis, rehabilitation, and low-vision ai</li> <li>8.2.1 Constrained text generation to meas</li> <li>8.2.2 Inclusive news design: enhancing resoptimization</li></ul>	ds       18         ure reading performance       18         ading comfort on small displays through layout       18			
	<ul> <li>8.3 Visual media analysis and creation</li> <li>8.3.1 Human Trajectory Forecasting in 3E Low Vision</li></ul>	20         Environments: Navigating Complexity under         20         erize immersive user experience with multivari-			
	8.3.4 HandyNotes: using the hands to cr aware real-world objects	eate semantic representations of contextually			
	8.3.5 AMD Journee: A patient co-designed impact of AMD on social interaction	I VR experience to raise awareness towards thes24			

9	Bilateral contracts and grants with industry	25
	9.1 Bilateral contracts with industry	25
10 Partnerships and cooperations 2		
	10.1 International initiatives10.1.1 Associate Teams in the framework of an Inria International Lab or in the framework	26
	of an Inria International Program	26
	10.1.2 Participation in other International Programs	26
	10.2 International research visitors	27
	10.2.1 Visits of international scientists	27
	10.3 National initiatives	27
	10.3.1 ANR	27
11	Dissemination	29
	11.1 Promoting scientific activities	29
	11.1.1 Scientific events: organisation	29
	11.1.2 Scientific events: selection	29
	11.1.3 Journal	30
	11.1.4 Invited talks	30
	11.1.5 Scientific expertise	30
	11.1.6 Research administration	30
	11.2 Teaching - Supervision - Juries	30
	11.2.1 Teaching	30
	11.2.2 Supervision	31
	11.2.3 Juries	32
	11.3 Popularization	32
	11.3.1 Other science outreach relevant activities	32
12	2 Scientific production	32
	12.1 Major publications	32
	12.2 Publications of the year	33
	12.3 Cited publications	35

# **Project-Team BIOVISION**

Creation of the Project-Team: 2018 August 01

# Keywords

# Computer sciences and digital sciences

- A5.1.1. Engineering of interactive systems
- A5.1.2. Evaluation of interactive systems
- A5.1.9. User and perceptual studies
- A5.3. Image processing and analysis
- A5.4. Computer vision
- A5.5.4. Animation
- A5.6.1. Virtual reality
- A5.6.2. Augmented reality
- A5.8. Natural language processing
- A6.1.1. Continuous Modeling (PDE, ODE)
- A6.1.4. Multiscale modeling
- A6.1.5. Multiphysics modeling
- A6.2.4. Statistical methods
- A6.3.3. Data processing
- A7.1.3. Graph algorithms
- A9.4. Natural language processing
- A9.7. AI algorithmics

## Other research topics and application domains

- B1.1.8. Mathematical biology
- B1.2. Neuroscience and cognitive science
- B1.2.1. Understanding and simulation of the brain and the nervous system
- B1.2.2. Cognitive science
- B1.2.3. Computational neurosciences
- B2.1. Well being
- B2.5.1. Sensorimotor disabilities
- B2.5.3. Assistance for elderly
- B2.7.2. Health monitoring systems
- B9.1.2. Serious games
- B9.3. Medias
- B9.5.2. Mathematics
- B9.5.3. Physics
- B9.6.8. Linguistics
- B9.9. Ethics

# 1 Team members, visitors, external collaborators

# **Research Scientists**

- Bruno Cessac [Team leader, INRIA, Senior Researcher, HDR]
- Pierre Kornprobst [INRIA, Senior Researcher, HDR]
- Hui-Yin Wu [INRIA, ISFP]

# **Post-Doctoral Fellow**

• Sebastian Santiago Vizcay [INRIA, Post-Doctoral Fellow, until Feb 2024]

# **PhD Students**

- Alexandre Bonlarron [INRIA, until Sep 2024]
- Johanna Delachambre [UNIV COTE D'AZUR]
- Pauline Devictor [UNIV COTE D'AZUR, from Nov 2024]
- Jerome Emonet [INRIA]
- Franz Franco Gallo [INRIA]
- Sebastian Gallardo Diaz [INRIA, CIFRE]
- Erwan Petit [UNIV COTE D'AZUR]
- Florent Robert [UNIV COTE D'AZUR, until Sep 2024]

# **Technical Staff**

• Clement Bergman [INRIA, Engineer, until May 2024]

# **Interns and Apprentices**

- Gillian Arlabosse [TELECOM SUDPARIS, Intern, from Jul 2024 until Aug 2024]
- Pauline Devictor [UNIV COTE D'AZUR, Intern, from Mar 2024 until Sep 2024]
- Christos Kyriazis [INRIA, Intern, from Mar 2024 until Aug 2024]
- Christos Kyriazis [UNIV COTE D'AZUR, Intern, until Feb 2024]
- Laura Piovano [UNIV COTE D'AZUR, Intern, from Sep 2024]
- Laura Piovano [UNIV COTE D'AZUR, Intern, from Mar 2024 until May 2024]

# Administrative Assistant

• Marie-Cecile Lafont [INRIA]

# **Visiting Scientists**

- Francisco Miqueles Varela [UNIV CHILI, from Mar 2024 until May 2024]
- Jorge Portal Diaz [UNIV CHILI, from Mar 2024 until May 2024]

## **External Collaborators**

- Aurélie Calabrese [CNRS]
- Eric Castet [CNRS]

# 2 Overall objectives

Vision is a key function to sense our world and perform complex tasks. It has high sensitivity and strong reliability, even though most of its input is noisy, changing, and ambiguous. A better understanding of how biological vision works opens up scientific challenges as well as promising technological, medical and societal breakthroughs. Fundamental aspects such as understanding how a visual scene is encoded by the retina into spike trains, transmitted to the visual cortex via the optic nerve through the thalamus, decoded in a fast and efficient way, and then creating a sense of perception, offers perspectives in research and technological developments for current and future generations.

Vision is not always functional though. Sometimes, "something" goes wrong. Although many visual impairments such as myopia, hypermetropia, cataract, can be cured by glasses, contact lenses, or other means like medicine or surgery, pathologies impairing the retina such as Age-Related Macular Degeneration (AMD) and Retinis Pigmentosa (RP) can not be fixed with these standard treatments [39]. They result in a progressive degradation of vision (Figure 1), up to a stage of low vision (visual acuity of less than 6/18 to light perception, or a visual field of less than 10 degrees from the point of fixation) or blindness. Thus, people with low vision must learn to adjust to their pathologies. Progress in research and technology can help them. Considering the aging of the population in developed countries and its strong correlation with the prevalence of eye diseases, low vision has already become a major societal problem.



Figure 1: Central blind spot (i.e., scotoma), as perceived by an individual suffering from Central Field Loss (CFL) when looking at someone's face.

In this context, the Biovision Team's research revolves around the central theme biological vision and perception, and the impact of low vision conditions. Our strategy is based upon four cornerstones: To model, to assist diagnosis, to aid visual activities like reading, and to enable personalized content creation. We aim to develop fundamental research as well as technology transfer along three entangled axes of research:

• Axis 1: Modeling the retina and the primary visual system.

- Axis 2: Diagnosis, rehabilitation, and low-vision aids.
- Axis 3: Visual media analysis and creation.

These axes form a stable, three-pillared basis for our research activities, giving our team an original combination in expertise: modeling for the neurosciences, computer vision, Virtual and Augmented Reality (XR), and media analysis and creation. Our research themes require strong interactions with experimental neuroscientists, modelers, ophtalmologists and patients, constituting a large network of national and international collaborators. Biovision is therefore a strongly multi-disciplinary team. We publish in international reviews and conferences in several fields including neuroscience, low vision, mathematics, physics, computer vision, multimedia, computer graphics, and human-computer interactions.

# **3** Research program

# 3.1 Axis 1 - Modeling the retina and the primary visual system.

In collaboration with neuroscience labs, we derive phenomenological equations and analyze them mathematically by adopting methods from theoretical physics or mathematics (Figure 2). We also develop simulation platforms like Pranas or Macular, helping us confront theoretical predictions to numerical simulations, or allowing researchers to perform *in silico* experimentation under conditions rarely accessible to experimentalists (such as simultaneously recording the retina layers and the primary visual cortex<sup>1</sup> (V1)). Specifically, our research focuses on the modeling and mathematical study of:

- Multi-scale dynamics of the retina in the presence of spatio-temporal stimuli;
- Response to motion, anticipation and surprise in the early visual system;
- Spatio-temporal coding and decoding of visual scenes by spike trains;
- Retinal pathologies.

#### 3.2 Axis 2 - Diagnosis, rehabilitation, and low-vision aids.

In collaboration with low vision clinical centers and cognitive science labs, we develop computer science methods, open software and toolboxes to assist low vision patients (Figure 3), with a particular focus on Age-Related Macular Degeneration<sup>2</sup>. As AMD patients still have a plastic and functional vision in their peripheral visual field [45], they must develop efficient "Eccentric Viewing" (EV) to adapt to the central blind zone (scotoma) and to direct gaze away from the object they want to identify [54]. Commonly proposed assistance tools involve visual rehabilitation methods [47] and visual aids that usually consist of magnifiers [40].

Our main research goals are:

- Understanding the relations between anatomo-functional and behavioral observations;
- Diagnosis from reading performance screening and oculomotor behavior analysis;
- Personalized and gamified rehabilitation for training eccentric viewing in VR;
- Personalized visual aid systems for daily living activities.

<sup>&</sup>lt;sup>1</sup>By "primary visual cortex" we mean V1. We use the terminology "primary visual system" or "early visual system" for the system retina, Lateral Geniculate Nucleus-LGN and V1.

<sup>&</sup>lt;sup>2</sup>This is the leading cause of irreversible loss of vision for subjects over the age of 65 in industrialized countries [60]. AMD patients who are blind in their central visual field and perceptual functions (notably reading and face/object identification) are dramatically impacted with loss of autonomy, degraded quality of life, and, often, depression [56, 42].



Figure 2: The process of retina modeling. A) The retina structure from biology. B) Designing a simplified architecture keeping retina components that we want to better understand (here, the role of Amacrine cells in motion anticipation); C) Deriving mathematical equations from A and B. D, E). Results from numerical modeling and mathematical modeling. Here, our retina model's response to a parabolic motion where the peak of response resulting from amacrine cells connectivity or gain control is in advance with respect to the peak in response without these mechanisms. This illustrates retinal anticipation.



Figure 3: Multiple methodologies in graphics and image processing have applications towards lowvision technologies including (a) 3D virtual environments for studies of user perception and behavior, and creating 3D stimuli for model testing, (b) image enhancement techniques to magnify and increase visibility of contours of objects and people, and (c) personalization of media content such as text in 360 degrees visual space using VR headsets.

# 3.3 Axis 3 - Visual media analysis and creation.

We investigate the impact of visual media design on user experience and perception, and propose assisted creativity tools for creating personalized and adapted media content (Figure 4). We employ computer vision and deep learning techniques for media understanding in film and in complex documents like newspapers. We deploy this understanding in new media platforms such as virtual and augmented reality for applications in low-vision training, accessible media design, and generation of 3D visual stimuli:

- Accessible / personalized media design for low vision training and reading platforms;
- Assisted creativity tools for 3D environments;
- Visual system and multivariate user behavior modeling in 3D contextual environments;
- Visual media understanding and gender representation in film.



Figure 4: We try to understand the perception of users in face of media and the gaps therein – on ontology, intersubjectivity, and intentionality – that result from digital interfaces that mediate the digital content and the communication between content designers and end users.

# 4 Application domains

# 4.1 Applications of low-vision studies

- **Cognitive research**: Virtual reality technology represents a new opportunity to conduct cognitive and behavioral research in virtual environments where all parameters can be psychophysically controlled. In the scope of ANR DEVISE, we are currently developing and using the PTVR software (Perception Toolbox for Virtual Reality) [16] to develop our own experimental protocols to study low vision. However, we believe that the potential of PTVR is much larger as it could be useful to any researcher familiar with Python programming willing to create and analyze a sophisticated experiment in VR with parsimonious code.
- Serious games: Serious games use game mechanics in order to achieve goals such as in training, education, or awareness. In our context, we want to explore serious games as a way to help low-vision patients in performing rehabilitation exercises. Virtual and augmented reality technology is a promising platform to develop such rehabilitation exercises targeted to specific pathologies due to their potential to create fully immersive environments, or inject additional information in the real world. For example, with Age-Related Macular Degeneration (AMD), our objective is to propose solutions allowing rehabilitation of visuo-perceptual-motor functions to optimally use residual portions of the peripheral retina defined from anatomo-functional exams [43].

• Vision aid-systems: A variety of aids for low-vision people are already on the market. They use various kinds of desktop (e.g., CCTVs), handheld (mobile applications), or wearable (e.g. OxSight, Helios) technologies, and offer different functionalities including magnification, image enhancement, text to speech, face and object recognition. Our goal is to design new solutions allowing autonomous interaction primarily using mixed reality – virtual and augmented reality. This technology could offer new affordable solutions developed in synergy with rehabilitation protocols to provide personalized adaptations and guidance.

# 4.2 Applications of vision modeling studies

• Neuroscience research. Making in-silico experiments is a way to reduce the experimental costs, to test hypotheses and design models, and to test algorithms. Our goal is to develop a large-scale simulations platform of the normal and impaired retinas. This platform, called Macular, allows to test hypotheses on the retina functions in normal vision (such as the role of amacrine cells in motion anticipation [58], or the expected effects of pharmacology on retina dynamics [44]). It is also to mimic specific degeneracies or pharmacologically induced impairments [53], as well as to emulate electric stimulation by prostheses.

In addition, the platform provides a realistic entry to models or simulators of the thalamus or the visual cortex, in contrast to the entries usually considered in modeling studies.

• Education. Macular is also targeted as a useful tool for educational purposes, illustrating for students how the retina works and responds to visual stimuli.

# 4.3 Applications of visual media analysis and creation

- Engineering immersive interactive storytelling platforms. With the sharp rise in popularity of immersive virtual and augmented reality technologies, it is important to acknowledge their strong potential to create engaging, embodied experiences. Reaching this goal involves investigating the impact of virtual and augmented reality platforms from an interactivity and immersivity perspective: how do people parse visual information, react, and respond in the face of content, in the face of media [15, 11, 50, 49].
- **Personalized content creation and assisted creativity.** Models of user perception can be integrated in tools for content creators, such as to simulate low-vision conditions to aid in the design of accessible spaces and media, and to diversify immersive scenarios used for studying user perception, for entertainment, and for rehabilitation [14, 61].
- Media studies and social awareness. Investigating interpretable models of media analysis will allow us to provide tools to conduct qualitative media studies on large amounts of data in relation to existing societal challenges and issues. This includes understanding the impact of media design on accessibility for patients with low-vision, as well as raising awareness towards biases in media [62] and existing media analysis tools from deep learning paradigms [63].

# 5 Social and environmental responsibility

The research themes in the Biovision team have direct social impacts on two fronts:

- Low vision: we work in partnership with neuroscientists and ophthalmologists to design technologies for the diagnosis and rehabilitation of low-vision pathologies, addressing a strong societal challenge.
- Accessibility: in concert with researchers in media studies, we tackle the social challenge of designing accessible media, including for the population with visual impairments, as well as to address media bias both in content design and in machine learning approaches.

# 6 Highlights of the year

# 6.1 Awards

- Simone Ebert got the first thesis prize of the EDSTIC, mention COP/ELEC/SN, for her PhD "Dynamical synapses in the retinal network" completed in 2023 under the supervision of B. Cessac.
- We published the paper "Temporal pattern recognition in retinal ganglion cells is mediated by dynamical inhibitory synapses" in Nature Communications.
- We obtained an ECOS-SUD funding on the project CANARD, in collaboration with the University of Valparaiso, Chile and the Federal University of Minas Gerais (Brasil) for the year 2025.
- Franz Franco Gallo received a best poster award at the SophIA Summit.

# 7 New software, platforms, open data

# 7.1 New software

#### 7.1.1 Macular

Name: Numerical platform for simulations of the primary visual system in normal and pathological conditions

#### Keywords: Retina, Vision, Neurosciences

**Scientific Description:** Macular is built around a central idea: its use and its graphical interface can evolve according to the user's objectives. It can therefore be used in user-designed scenarios, such as simulation of retinal waves, simulation of retinal and cortical responses to prosthetic stimulation, study of pharmacological impact on retinal response, etc. The user can design their own scenarios using the Macular Template Engine.

At the heart of Macular is an object called "Cell". Basically these "cells" are inspired by biological cells, but it's more general than that. It can also be a group of cells of the same type, a field generated by a large number of cells (for example a cortical column), or an electrode in a retinal prosthesis. A cell is defined by internal variables (evolving over time), internal parameters (adjusted by cursors), a dynamic evolution (described by a set of differential equations) and inputs. Inputs can come from an external visual scene or from other synaptically connected cells. Synapses are also Macular objects defined by specific variables, parameters, and equations. Cells of the same type are connected in layers according to a graph with a specific type of synapses (intra-layer connectivity). Cells of a different type can also be connected via synapses (inter-layer connectivity).

All the information concerning the types of cells, their inputs, their synapses and the organization of the layers are stored in a file of type .mac (for "macular") defining what we call a "scenario". Different types of scenarios are offered to the user, which they can load and play, while modifying the parameters and viewing the variables (see technical section).

**Functional Description:** Macular is a simulation platform for the retina and the primary visual cortex, designed to reproduce the response to visual stimuli or to electrical stimuli, in normal vision conditions, or altered (pharmacology, pathology, development).

#### Release Contributions: First release.

#### URL: https://team.inria.fr/biovision/macular-software/

#### Contact: Bruno Cessac

Participants: Bruno Cessac, Evgenia Kartsaki, Selma Souihel, Jerome Emonet, Sebastian Gallardo Diaz, Ghada Bahloul, Tristan Cabel, Erwan Demairy, Pierre Fernique, Thibaud Kloczko, Come Le Breton, Jonathan Levy, Nicolas Niclausse, Jean-Luc Szpyrka, Julien Wintz, Carlos Zubiaga Pena

#### 7.1.2 GUsT-3D

Name: Guided User Tasks Unity plugin for 3D virtual reality environments

Keywords: 3D, Virtual reality, Interactive Scenarios, Ontologies, User study

**Functional Description:** We present the GUsT-3D framework for designing Guided User Tasks in embodied VR experiences, i.e., tasks that require the user to carry out a series of interactions guided by the constraints of the 3D scene. GUsT-3D is implemented as a set of tools that support a 4-step workflow to : (1) annotate entities in the scene with names, navigation, and interaction possibilities, (2) define user tasks with interactive and timing constraints, (3) manage scene changes, task progress, and user behavior logging in real-time, and (4) conduct post-scenario analysis through spatio-temporal queries on user logs, and visualizing scene entity relations through a scene graph.

The software also includes a set of tools for processing gaze tracking data, including: cleaning and synchronization of the data, calculation of fixations with I-VT, I-DT, IDTVR, IS5T, Remodnav, and IDVT algorithms, and visualization of the data (points of regard and fixations) in both real time and collectively.

**News of the Year:** A new version of the software has been released with additional functionalities for collection, processing, and analysis of gaze tracking. We also made major changes to the overall workflow (system logging, interactions, bug fixes, etc.). This version was used in two user studies.

URL: https://project.inria.fr/creattive3d/gust-3d/

Publications: hal-04102737, hal-04446066, hal-03635452

Contact: Hui-Yin Wu

Participants: Hui-Yin Wu, Marco Alba Winckler, Lucile Sassatelli, Florent Robert

Partner: I3S

#### 7.1.3 PTVR

Name: Perception Toolbox for Virtual Reality: The Open-Source Python Library for Virtual Reality Experiments

Keywords: Visual perception, Behavioral science, Virtual reality

**Functional Description:** PTVR is a fully open-source library for creating visual perception experiments in virtual reality using a high-level Python interface. It revolutionizes the way you design and conduct VR experiments, combining power and simplicity to bring your ideas to life:

(1) Effortless 3D Placement: With our innovative "Perimetric" coordinate system, positioning stimuli in 3D space is as simple as it gets. (2) Visual Angle Precision: Easily manage visual angles in a three-dimensional environment for precise experimental control. (3) Flat Screen Simulation: Replicate and extend standard 2D screen experiments on virtual flat screens, seamlessly bridging the gap between 2D and VR. (4) Data at Your Fingertips: Save experimental results and capture behavioral data, including head and gaze tracking, with ease. (5) Interactive Documentation: Our user-friendly guides include animated 3D figures that demystify the subtleties of working in a 3D space. (6) Hands-On Learning: Dive into the action with a suite of demo scripts, designed to be experienced directly through a VR headset—perfect for quick mastery of PTVR's capabilities.

New for 2024: Import Realistic 3D Assets into Your Experiments. PTVR now supports the integration of complex and realistic 3D objects—commonly referred to as "assets" in the VR world. From furniture to detailed animals, characters, and even entire cities, you can leverage pre-made assets available for purchase or free download. With Unity's "asset bundle" system, importing and managing large collections of assets is a breeze. PTVR's intuitive syntax allows you to incorporate these assets directly into your experiments, expanding your creative possibilities. Whether you're a researcher, educator, or VR enthusiast, PTVR is your go-to Python tool for crafting innovative, immersive, and precise virtual reality experiments. Join the open-source community today and take your experiments to the next dimension!

URL: https://ptvr.inria.fr

Publication: hal-04579084

Contact: Pierre Kornprobst

**Partner:** Aix-Marseille Université - CNRS Laboratoire de Psychologie Cognitive - UMR 7290 - Team 'Perception and attention'

#### 7.1.4 Tatoovi

Name: Tagging tool for videos

Keywords: Annotation tool, Video analysis, Multimedia player, Data visualization

**Functional Description:** Tatoovi is an application which provides different modules for video annotation including: (1) a customizable annotation dictionary with visual editor that includes fields for classes, numerical values and scales, labelling, and text, (2) multi-level annotation by keyframes, shots/clips, and interface for film metadata and character information, (3) a click-and-drag pose module to edit and visualize character skeletons and bounding boxes, (4) and multi-timeline visualization and intuitive video playback and navigation tools, and (5) separate JSON export of timelines, annotations, and dictionaries for ease of collaboration, data analysis, and machine learning models

Contact: Hui-Yin Wu

Participants: Hui-Yin Wu, Lucile Sassatelli, Clement Bergman, Genevieve Masioni Kibadi, Luan Nguyen

Partner: I3S

# 7.2 New platforms

Members of Biovision are marked with a  $\star$ .

#### 7.2.1 CREATTIVE3D platform for navigation studies in VR

**Participants:** ★ Hui-Yin Wu, ★ Florent Robert, Lucile Sassatelli (*UniCA, CNRS, I3S; Institut Universitaire de France*), Marco Winckler (*UniCA, CNRS, I3S*).

As part of ANR CREATTIVE3D, the Biovision team has established a technological platform in the Kahn immersive space including:

- a  $40m^2$  tracked space for the Vive Pro Eye virtual reality headset with 4-8 infra-red base stations,
- GUsT-3D software 7.1.2 under Unity for the creation, management, logging, analysis, and scene graph visualization of interactive immersive experiences,
- Vive Pro Eye headset integrated sensors including gyroscope and accelerometers, spatial tracking through base stations, and 120 Hz eye tracker,
- External sensors including XSens Awinda Starter inertia-based motion capture costumes and Shimmer GSR physiological sensors.

The platform has hosted around 60 user studies lasting over 120 hours this year, partly published in [11]. It has also provided demonstrations of our virtual reality projects for visitors, collaborators, and students.

# 7.3 Open data

# CREATTIVE3D multimodal dataset of user behavior in virtual reality

#### **Contributors:**

**Participants:** \* Hui-Yin Wu, \* Florent Robert, Lucile Sassatelli (UniCA, CNRS, I3S), Marco Winckler (UniCA, Polytech, CNRS, I3S), Stephen Ramanoël (UniCA, LAMHESS), Auriane Gros (UniCA, CHU Nice, CoBTeK).

**Description:** In the context of the ANR CREATTIVE3D project, we present the CREATTIVE3D dataset of human interaction and navigation at road crossings in virtual reality. It is the largest dataset of human motion in fully-annotated scenarios (40 hours, 2.6 million poses), it captured in dynamic 3D scenes with multivariate – gaze, physiology, and motion – data. Its purpose is to investigate the impact of simulated low-vision conditions using dynamic eye tracking under real walking and simulated walking conditions. Extensive effort has been made to ensure the transparency, usability, and reproducibility of the study and collected data, even under extremely complex study conditions involving 6 degrees of freedom interactions, and multiple sensors.

Dataset PID (DOI): https://doi.org/10.5281/zenodo.8269108

Project link: https://project.inria.fr/creattive3d/

**Publications:** [11, 12, 7]

Contact: Hui-Yin Wu

#### Visual Objectification in Films: Towards a New AI Task for Video Interpretation

#### **Contributors:**

Participants: Julie Tores (UniCA, CNRS, I3S), Lucile Sassatelli (UniCA, CNRS, I3S), \* Hui-Yin Wu, \* Clément Bergman, Léa Andolfi (UniCA, CNRS, I3S), Victor Ecrement (Sorbonne Université, GRIPIC), Frédéric Precioso (UniCA, Inria, CNRS, I3S), Thierry Devars (Sorbonne Université, GRIPIC), Magali Guaresi (UniCA, CNRS, BCL), Virginie Julliard (Sorbonne Université, GRIPIC), Sarah Lecossais (Université Sorbonne Paris Nord, LabSIC).

**Description:** This is the first dataset for visual objectification in videos, specifically in films. We name this dataset ObyGaze12, short for Objectify-ingGaze12, which has the following highlights:

- It considers the multiple dimensions of the construct of visual objectification, made of filmic (framing and editing over successive shots, camera motion, etc.) and icono-graphic properties (visible objects, body parts, attire, character interactions, etc.).
- It is based on a thesaurus articulating five sub-constructs identified from multidisciplinary literature (film studies, psychology) from which we define typical instances, then grouped into coarse-grained visual concepts.
- The data is annotated densely with concepts, and shows the multi-factorial property of objectification, corroborating with some recent developments in cognitive psychology.
- Categories to annotate include a hard negative category meant to perform fine-grained error analysis and improve model generalization.

Dataset Link: https://github.com/husky-helen/ObyGaze12

Project link: https://webcms.i3s.unice.fr/TRACTIVE/home

Publications: [25]

Contact: Hui-Yin Wu

# 8 New results

We present here the new scientific results of the team over the course of the year. For each entry, members of Biovision are marked with a  $\star$ .

# 8.1 Modeling the retina and the primary visual system

# 8.1.1 Temporal pattern recognition in retinal ganglion cells is mediated by dynamical inhibitory synapses

Participants:\* Simone Ebert, Thomas Buffet (Institut de la Vision, Sorbonne Université, Paris, France.), Semihcan Sermet (Institut de la Vision, Sorbonne Université, Paris, France.), Olivier Marre (Institut de la Vision, Sorbonne Université, Paris, France.), \* Bruno Cessac.

**Description**: A fundamental task for the brain is to generate predictions of future sensory inputs, and signal errors in these predictions. Many neurons have been shown to signal omitted stimuli during periodic stimulation, even in the retina. However, the mechanisms of this error signaling are unclear. Here we show that depressing inhibitory synapses enable the retina to signal an omitted stimulus in a flash sequence. While ganglion cells, the retinal output, responded to an omitted flash with a constant latency over many frequencies of the flash sequence, we found that this was not the case once inhibition was blocked. We built a simple circuit model and showed that depressing inhibitory synapses was a necessary component to reproduce our experimental findings (Figure 5). We also generated new predictions with this model, which we confirmed experimentally. Depressing inhibitory synapses could thus be a key component to generate the predictive responses observed in many brain areas.

This work has been published in the journal Nature Communication [17].

# 8.1.2 How does the inner retinal network shape the ganglion cells receptive field : a computational study

Participants:\* Evgenia Kartsaki (Biosciences Institute, Newcastle University, Newcastle upon Tyne, UK; UniCA, Inria, SED), Gerrit Hilgen (Biosciences Institute, Newcastle University, Newcastle upon Tyne, UK; Health & Life Sciences, Applied Sciences, Northumbria University, Newcastle upon Tyne UK), Evelyne Sernagor (Biosciences Institute, Newcastle University, Newcastle

**Description**: We consider a model of basic inner retinal connectivity where bipolar (BCs) and amacrine (ACs) cells interconnect, and both cell types project onto ganglion cells (RGCs), modulating their response output to the brain visual areas. We derive an analytical formula for the spatio-temporal response of retinal ganglion cells to stimuli taking into account the effects of amacrine cells inhibition. This analysis reveals two important functional parameters of the network: (i) the intensity of the interactions between bipolar and amacrine cells, and, (ii) the characteristic time scale of these responses. Both parameters have a profound combined impact on the spatiotemporal features of RGC responses to light. The validity of the model is confirmed by faithfully reproducing pharmacogenetic experimental results obtained by stimulating excitatory DREADDs (Designer Receptors Exclusively Activated by Designer Drugs) expressed on ganglion cells and amacrine cells subclasses, thereby modifying the inner retinal network activity to visual stimuli in a complex, entangled manner. DREADDs are activated by "designer drugs" such as clozapine-n-oxide (CNO), resulting in an increase in free cytoplasmic calcium and profound increase in excitability in the cells that express these DREADDs (Roth, 2016). We found DREADD expression both in subsets of RGCs and in ACs in the inner nuclear layer (INL) [51]. In these conditions, CNO would act both



Figure 5: Glycinergic Amacrine cells are necessary for predictive timing of the OSR. **A.** Schematic representation of the retina, with the activity of retinal ganglion cells being recorded with a multi electrode array. **B.** Temporal traces of the receptive fields of the cells that exhibit an Omitted Stimulus Response (OSR). **C.** Example of OSR. The cell responds to the end on the stimulation, after the last flash. The dotted rectangle shows the time period of focus in panel D. **D.** Experimental recording of the OSR in one cell in control condition (black) and with strychnine to block glycinergic amacrine cell transmission (red). Firing rate responses to flash trains of 3 different frequencies are aligned to the last flash of each sequence. Flashes are represented by gray patches, vertical lines indicate the maximum of the response peak, red shaded areas indicate the temporal discrepancy between control and strychnine conditions.

on ACs and RGCs providing a way to unravel the entangled effects of: (i) direct excitation of DREADDexpressing RGCs and increase inhibitory input onto RGCs originating from DREADD-expressing ACs; (ii) change in cells response timescale (via a change in the membrane conductance), thereby providing an experimental set up to validate our theoretical predictions. Our mathematical model allows us to explore and decipher these complex effects in a manner that would not be feasible experimentally and provides novel insights in retinal dynamics.

Figure 6 illustrates our results. This paper has been published in the journal *Neural Computation* [18].

#### 8.1.3 A chimera model for motion anticipation in the retina and the primary visual cortex

Participants:\* Jérôme Emonet, Selma Souihel (Inria, P16 - Programme IA), \*<br/>Bruno Cessac, Alain Destexhe (NeuroPSI - Institut des Neurosciences<br/>Paris-Saclay), Frédéric Chavane (INT - Institut de Neurosciences de la<br/>Timone), Matteo Di Volo (UCBL - Université Claude Bernard Lyon 1).

Description: Motion anticipation refers to the capacity of the visual system to compensate for inherent delays in visual processing. This ability results from distinct mechanisms taking place in the retina [38] and in the visual cortex [52]. To study their respective role, we propose a mean field model of the primary visual cortex (V1) connected to a realistic retina model. Our first goal is to reproduce experimental results on motion anticipation, made in monkeys by using voltage dye optical imaging (VSDI) [37], and to assess the impact of the retina in this process. For this, we first study the model in the case where the retina does not itself provide anticipation. Then, anticipation is only triggered by a cortical mechanism, called "anticipation by latency". As we show, this mechanism strongly depends on the intensity of the retinal input supplied to the cortex, even if the retina does not itself provide anticipation. We also unravel the effect of the stimulus features, such as speed and contrast, and report the impact of physiological parameters not accessible experimentally, such as the size of cortical extensions or fibre conduction. Then, we explore the changes in the cortical wave of anticipation when V1 is triggered by a retina output implementing different potential retina-driven anticipatory mechanisms, including gain control and lateral inhibition by amacrine cells. In this setting, we show how retinal and cortical anticipation combine, to provide an efficient processing where the VSDI signal response is in advance over the moving object that triggers this response, compensating the delays in visual processing, in full agreement with the experimental results of [37]. This work has been presented in the conferences "Annual Neuromod Meeting" [26], "AREADNE 2024" [33], ICMNS 2024 [32] and submitted to the journal Neural Computation. An example of results is shown in Fig. 7

#### 8.1.4 Optical Flow Estimation pre-training with simulated stage II retinal waves

**Participants:** \* Christos Kyriazis, \* Bruno Cessac, \* Hui-Yin Wu, \* Pierre Kornprobst.

**Context**: Retinal waves are an early developmental process in the mammalian visual system, playing a crucial role in structuring retinal connectivity and priming the system for visual functions such as motion detection. In [30], we investigated innovative approaches to leveraging simulated retinal waves (RWs) as a form of pre-training for machine learning (ML) models, particularly in the domain of optical flow estimation (OFE).

**Description**: We focus on a recent non-transformer machine learning architecture, RAFT [59], and employ a transfer learning strategy to harness the potential of retinal waves (RWs) in improving optical flow estimation (OFE) performance through a related task. Additionally, we explore an alternative approach that approximates the optical flow of RWs, enabling their direct application within OFE. While the concept of leveraging these biologically inspired stimuli to create more accessible training data holds promise for enhancing the generalization capabilities of OFE models, our findings indicate limited effectiveness with both approaches.



Figure 6: CNO may change the cell response. The upper left panels shows the representative point of a cell, labelled  $83_2$ , in a two dimensional map, called the RF map, in control (CTL - purple point) and CNO (red point) conditions. This shows the displacement of the representative point of the cell when CNO is applied. The **upper right panel** shows the same motion, in the same space (*r*, *s*), but considering here the main peak in the power spectrum, corresponding to a characteristic period  $T_1$ . The dashed black lines correspond to the separation between the monophasic, biphasic, and polyphasic regions. One switches from a region with a high  $T_1$  of order 600 ms (green area) in CTL, to a region where  $T_1$  is of order 300 ms (grey area) in CNO conditions. As shown in the **bottom panel**, this displacement corresponds to a switch from monophasic to biphasic region.



Figure 7: The effect of bipolar gain control strength,  $h_B$ , on the cortical response. (see the paper [32] for detail.) VSDI signal response to bipolar gain control at A) 0 mV/s (control) and B) 9.2 mV/s. C) Temporal VSDI signal in response to increasing bipolar gain control for the cortical column located at the center of the layer ( $x = 9^\circ$ ,  $y = 1.35^\circ$ ). D) Spatial VSDI signal in response to increasing bipolar gain control, for the time where the central cortical column reaches its maximum. The x coordinates has been shifted so that the central cortical column is actually located at x = 0. E) VSDI signal amplitude of the central cortical column versus  $h_B$ . F) Temporal and spatial observables: anticipation range (green) and maximal latency (red) versus  $h_B$ . G) Speed observable : short-range activation speed (red) in function of bipolar gain control weight. In A, B, C, we also drawn the Equivalent Retinal Output Amplitude (EROA) curve. This is the dotted black curve with the same colored symbols. H) Stationary peak delay (SPD) for RGCs (orange), VSDI signal (blue) for the central cell (scales on the left) and difference between RGC SPD and VSDI signal SPD (black, scales on the right). I) Shape of the central RGC response profile to the moving bar, without gain control (red) and with BC gain control 9.165 mV/s (dashed green). Note that the two traces have been rescaled to have the same maximum. This is to emphasize the change in the shape of the response induced by BCs gain control. J) VSDI signal, same conditions. In I,J, the dotted lines correspond to the peaks in the RGC firing rate or VSDI signal without BCs gain control (red) and with it (green).

## 8.2 Diagnosis, rehabilitation, and low-vision aids

#### 8.2.1 Constrained text generation to measure reading performance

Participants:\* Alexandre Bonlarron, Aurélie Calabrèse (Aix-Marseille Université<br/>(CNRS, Laboratoire de Psychologie Cognitive, Marseille, France)), \*<br/>Pierre Kornprobst, Jean-Charles Régin (UniCA, I3S, Constraints and<br/>Application Lab).

**Context**: Measuring reading performance is one of the most widely used methods in ophthalmology clinics to judge the effectiveness of treatments, surgical procedures, or rehabilitation techniques [57]. However, reading tests are limited by the small number of standardized texts available. For the MNREAD test [55], which is one of the reference tests used as an example in this work, there are only two sets of 19 sentences in French. These sentences are challenging to write because they have to respect rules of different kinds (e.g., related to grammar, length, lexicon, and display). They are also tricky to find: out of a sample of more than three million sentences from children's literature, only four satisfy the criteria of the MNREAD reading test.

**Description**: To obtain more sentences, we proposed in [41] a new standardized sentence generation method. Using this approach, we generated under constraints all possible MNREAD-like sentences that can be built from a given corpus and ranked them from 'good' to 'poor' using a transformers-based language model (GPT-2). To validate our method, we asked 14 normally sighted participants (age 14 to 56) to read 3 sets of 30 French sentences: MNREAD sentences, generated sentences ranked as 'good' and generated sentences ranked as 'poor'. All were displayed at 40cm, in regular polarity with a fixed print size of 1.3 logMAR. Corrected reading speed was measured in words/min (wpm) and analyzed using a mixed-effects model. Results were presented at ARVO 2024 conference [31]. On average, 'good' generated sentences were read at 164 wpm. This value was not significantly different from the reading speed of 162 wpm yield by MNREAD sentences (p=0.5). On the other hand, reading speed was significantly slower for 'poor' generated sentences, with an average value of 151 wpm (p<0.001). Our method seems to provide valid standardized sentences that follow the MNREAD rules and yield similar performance, at least in French. Since our method is easily applicable to other languages, further investigations is needed to validate the generation of sentences in English, Spanish, Italian, etc.

## 8.2.2 Inclusive news design: enhancing reading comfort on small displays through layout optimization

**Participants:** \* Sebastian Gallardo, Dorian Mazauric (*UniCA, Inria, ABS Team*), \* Pierre Kornprobst.

**Context**: The digital era transforms the newspaper industry, offering new digital user experiences for all readers. However, to be successful, newspaper designers stand before a tricky design challenge: translating the design and aesthetics from the printed edition (which remains a reference for many readers) and the functionalities from the online edition (continuous updates, responsiveness) to create the e-newspaper of the future, making a synthesis based on usability, reading comfort, engagement. In this spirit, our project aims to develop a novel inclusive digital news reading experience that will benefit all readers: you, me, and low vision people for whom newspapers are a way to be part of a well-evolved society.

**Description**: In this work we focus on how to comfortably read newspapers on a small display. Simply transposing the print newspapers into digital media can not be satisfactory because they were not designed for small displays. One key feature lost is the notion of entry points that are essential for navigation. By focusing on headlines as entry points, we show how to produce alternative layouts for small displays that preserve entry points quality (readability and usability) while optimizing aesthetics and style. Our approach consists in a relayouting approach implemented via a genetic-inspired approach. We tested it on realistic newspaper pages. For the case discussed here, we obtained more than 2000

different layouts where the font was increased by a factor of two. We show that the quality of headlines is globally much better with the new layouts than with the original layout (Figure 8). Future work will tend to generalize this promising approach, accounting for the complexity of real newspapers, with user experience quality as the primary goal.

More details are available in [48]. This work is still ongoing.



Figure 8: From print to online newspapers on small displays: (a) Print newspaper page where each article has been colored to highlight the structure. On small displays, headlines become too small to be readable and usable for navigation. Magnification is needed. (b) Pinch-zoom result with a magnification factor of two: Illustrates the common local/global navigation difficulty encountered when reading newspapers via digital kiosk applications. (c) Increasing the font with a factor two, keeping the original layout: Articles with unwanted shapes (denoted by UW, here when headlines flow on more than three lines) are boxed in red; (d) Best result of our approach showing how the page in (a) has been transformed to preserve headlines quality.

#### 8.2.3 A clustering based article template recommendation system for newspaper editors

 
 Participants:
 \*
 Sebastian Gallardo, Bruno Génuit (Demain un Autre Jour), Dorian Mazauric (UniCA, Inria, ABS Team), \*
 Pierre Kornprobst.

**Context**: Newspaper editors face significant challenges in page production, as printed newspapers are becoming less profitable and more costly to produce. Improving productivity in this process is crucial. In [29], we focus on the task of selecting the most appropriate article template for given content-a process that is time-consuming and difficult. We demonstrate how techniques from Recommendation Systems (RS) can be adapted and extended to assist newspaper editors by recommending the best templates based on their needs and preferences. We propose a clustering-based recommendation system that promotes diversity, which is a critical requirement in this context.

**Description**: Our method consists of two phases: first, an offline clustering phase that uses a graphmatching neural network to compute a custom similarity measure (GMN) between templates, modeled as graphs. This clustering process is based on content and structural information, independent of user preferences. The second phase is an online recommendation system that incorporates both user preferences and content requirements, promoting templates that minimize similarity. Compared to existing methods, our approach outperforms in terms of novelty and diversity, especially in cases where high-rated items are clustered together. We present promising results through case examples, demonstrating the practical applicability of our method for real-world scenarios, while also opening avenues for future studies with industry professionals. While this study provides a first step towards applying recommendation systems to optimize the selection of article templates in the newspaper production process, further research is needed to extend this work to full-page layouts and multi-page designs, incorporating more comprehensive recommendations that address the broader scope of newspaper production.

#### 8.3 Visual media analysis and creation

#### 8.3.1 Human Trajectory Forecasting in 3D Environments: Navigating Complexity under Low Vision

**Participants:** \* Franz Franco Gallo, \* Hui-Yin Wu, Lucile Sassatelli (UniCA, CNRS, I3S).

**Context:** This work tackles the challenge of predicting human trajectories while carrying out complex tasks in contextually-rich virtual environments. We evaluate the CREATIVE3D multimodal dataset on human interaction and navigation in 3D virtual reality (VR). In the dataset, navigating traffic crossings with simulated visual impairments are used as an example of complex or unpredictable situations. **Description:** In the first part of the work, we establish evaluations for a base multi-layer perceptron (MLP) and two state-of-the-art models: TRACK (RNN) and GIMO (transformer), on tasks with varying levels of complexity and visual impairment conditions. Our findings indicate that a model trained on normal visual conditions and simple tasks does not generalize on test data with complex interactions and simulated visual impairments, despite including 3D scene context and user gaze. In comparison, a model trained on diverse visual and task conditions is more robust, with up to 84% decrease in positional error and 9% in orientation error, but with the trade-off of lower accuracy for simpler tasks. We believe this work can benefit real-world applications such as autonomous driving, and enable context-aware computing for diverse scenarios and populations.



Figure 9: We use contextual graphs to help in a trajectory prediction task in virtual environments. The trajectories on the right compare two approaches to context inclusion: scene point clouds and scene graphs. Our model using scene graphs outperforms the model using scene point clouds.

In the second part of the work, we investigate how virtual environments provide a rich and controlled setting for collecting detailed data on human behavior, offering unique opportunities for predicting human trajectories in dynamic scenes. Most existing approaches have overlooked the potential of these environments, focusing instead on static contexts without considering user-specific factors. We propose Diverse Context VR Human Motion Prediction (DiVR), a cross-modal transformer based on the Perceiver architecture that integrates both static and dynamic scene context using a heterogeneous graph convolution network (Figure 10). We conduct extensive experiments comparing DiVR against existing architectures including MLP, LSTM, and transformers with gaze and point cloud context. Additionally, we

also test our model's generalizability across different users, tasks, and scenes. Results show that DiVR achieves higher accuracy and adaptability compared to other models and to static graphs. Some results are illustrated in Figure 9. This work highlights the advantages of using VR datasets for context-aware human trajectory modeling, with potential applications in enhancing user experiences in the metaverse. Our source code is publicly available on Inria Gitlab.



Figure 10: The DiVR model that adopts a perceiver architecture. The model takes as input the gaze context (gaze points interpolated with the scene point cloud), past motion, and the heterogeneous scene graph, and outputs a prediction on the future trajectory.

These works were presented at the ACM MMSys Workshop on Immersive Mixed and Virtual Environment Systems [7] and The ECCV Workshop on Computer Vision for Metaverse [21, 36]. It received a best poster award at the SophIA summit.

#### 8.3.2 Task-based methodology to characterize immersive user experience with multivariate data

Participants:\* Florent Robert, \* Hui-Yin Wu, Marco Winckler (UniCA, Polytech,<br/>CNRS, I3S), Lucile Sassatelli (UniCA, CNRS, I3S).

**Context:** Virtual Reality (VR) technologies enable strong emotions compared to traditional media, stimulating the brain in ways comparable to real-life interactions. This makes VR systems promising for research and applications in training or rehabilitation, to imitate realistic situations. Nonetheless, the evaluation of the user experience in immersive environments is daunting, the richness of the media presents challenges to synchronize context with behavioral metrics in order to provide fine-grained personalized feedback or performance evaluation. The variety of scenarios and interaction modalities multiply this difficulty of user understanding in the face of lifelike training scenarios, complex interactions, and rich context.

**Description:** We propose a task-based methodology that provides fine-grained descriptions and analysis of the experiential user experience (UX) in VR that (1) aligns low-level tasks (i.e. take an object, go somewhere) with multivariate behavior metrics: gaze, motion, skin conductance (Figure 11), (2) defines performance components (i.e., attention, decision, and efficiency) and baselines to evaluate task performance (e.g., the time it takes for a user to carry out the task, whether they direct their gaze towards key elements, etc.), and (3) characterizes task performance with multivariate user behavior data. To illustrate our approach, we apply the task-based methodology to an existing dataset from a road crossing study in VR (task model shown in Figure 12). We find that the task-based methodology allows us to better observe the experiential UX by highlighting fine-grained relations between behavior profiles and task performance, opening pathways to personalized feedback and experiences in future VR applications. This work was published at the 2024 IEEE Conference on Virtual Reality [12].



Figure 11: An example of the raw multivariate data (i.e., skin conductance (EDA), point of regard positions with x, y, z axis (POR), chest rotation value) synchronized with the context for one user in a single scenario. During this scenario, the users carried out four tasks: (1) Go outside their house, (2) Cross the road safely, (3) Take the box, (4) Cross the road safely to come back home.

#### 8.3.3 Visual Objectification in Films: Towards a New AI Task for Video Interpretation

 Participants: Julie Tores (UniCA, CNRS, I3S), Lucile Sassatelli (UniCA, CNRS, I3S), \* Hui-Yin Wu, \* Clément Bergman, Léa Andolfi (UniCA, CNRS, I3S), Victor Ecrement (Sorbonne Université, GRIPIC), Frédéric Precioso (UniCA, Inria, CNRS, I3S), Thierry Devars (Sorbonne Université, GRIPIC), Magali Guaresi (UniCA, CNRS, BCL), Virginie Julliard (Sorbonne Université, GRIPIC), Sarah Lecossais (Université Sorbonne Paris Nord, LabSIC).

**Description:** In film gender studies, the concept of "male gaze" refers to the way the characters are portrayed on the screen as objects of desire rather than subjects. In this work, we introduce a novel video-interpretation task, to detect character objectification in films. The purpose is to reveal and quantify the usage of complex temporal patterns operated in cinema to produce the cognitive perception of objectification. We introduce the ObyGaze12 dataset, made of 1914 movie clips densely annotated by experts for objectification concepts identified in film studies and psychology. Examples of films and annotations can be seen in Figure 13. We evaluate recent vision models, show the feasibility of the task with concept bottleneck models, and discuss the limitations. Our new dataset and code are made available to the community.

This work was presented as a poster highlight (top 11.9% of accepted papers) at the 2024 CVF Conference on Computer Vision and Pattern Recognition (CVPR) [25].

#### 8.3.4 HandyNotes: using the hands to create semantic representations of contextually aware realworld objects

Participants:	Clément Quéré (UniCA, Polytech, CNRS, I3S), Aline Menin (UniCA,	
	CNRS, I3S), Raphaël Julien (UniCA, Polytech), * Hui-Yin Wu,	
	Marco Winckler (UniCA, Polytech, CNRS, I3S).	

Description: We investigate Mixed Reality (MR) technologies to provide a seamless integration of digital



Figure 12: A simplified task model for three high level tasks: (a) "Go outside your house", (b) "Take the box", (c) "Cross the road safely", to properly achieve the task, the user is expected to perform the right interactions at the right time, and in the case of crossing the road, also look at traffic light and cars, and cross when the light is green and the cars are stopped. Rectangles indicate tasks and ovals indicate object states.



full body  $\rightarrow$  approach  $\rightarrow$  close-up on hips

Figure 13: In modern film media, the unequal characterization of gender on screen frequently evokes concepts of objectification, such as (A) unequal gaze (*Pulp Fiction*, 1994), (B) nudity and submissive postures (*Pulp Fiction*, 1994), (C) animalization or infantilization (*Marley and Me*, 2008), and (D) transparent clothing, camera framing, domestic gender roles, and voyeurism (*Gone Girl*, 2014).

information in physical environments through human-made annotations. Creating digital annotations of physical objects evokes many challenges for performing (simple) tasks such as adding digital notes and connecting them to real-world objects. For that, we have developed an MR system using the Microsoft HoloLens2 to create semantic representations of contextually-aware real-world objects while interacting with holographic virtual objects. User interaction is enhanced with use of fingers as placeholders for menu items. A figure of our application and its usage is shown in Figure 14. We demonstrate our approach through two real-world scenarios. We also discuss the challenges for using MR technologies.



Figure 14: Overview of the HandyNotes system. (a) Creation and manipulation of a bounding box determining the annotation target.(b) Annotation body associated with a bounding box. (c) Visualization of the object's annotations and the Hand-Menu System.

This work was published and presented at the 2024 IEEE Conference on Virtual Reality [23].

# 8.3.5 AMD Journee: A patient co-designed VR experience to raise awareness towards the impact of AMD on social interactions

Participants:	* Johanna Delachambre, Monica Di Meo (CHU - Hôpital Pasteur, Nice),	
	Frédérique Lagniez (CHU - Hôpital Pasteur, Nice), Christine Morfin-	
	Bourlat (CHU - Hôpital Pasteur, Nice), Stéphanie Baillif (CHU - Hôpital	
	Pasteur, Nice), * Hui-Yin Wu, * Pierre Kornprobst.	

**Context:** Virtual reality (VR) has great potential for raising awareness, thanks to its ability to immerse users in situations different from their everyday experiences, such as various disabilities. In [20], we present

a VR experience designed to raise awareness towards the impact of age-related macular degeneration (AMD) on social interactions for patients.

**Description:** We investigated age-related macular degeneration (AMD), which causes loss of central visual field acuity (a scotoma) and hinders patients from perceiving facial expressions and gestures. This often leads to social challenges such as awkward interactions, misunderstandings, and feelings of isolation. Using virtual reality (VR), we co-designed an immersive experience with four daily-life scenarios of AMD patients, informed by structured interviews with patients and orthoptists. The experience adopts the patient's perspective, incorporating voiceovers to narrate their emotions, challenges, adaptive strategies, and advice on how supportive social actions can improve their quality of life. A virtual scotoma, designed to follow the user's gaze via an HTC Vive Focus 3 headset with eye-tracking, enhances realism. Guided by a formal definition of awareness, we evaluate our experience on three components of awareness – knowledge, engagement, and empathy – through established questionnaires, continuous measures of gaze and skin conductance, and qualitative feedback. In an experiment with 29 participants, the experience demonstrated a strong positive impact on awareness about AMD. The scotoma and events also influenced gaze activity and emotional responses. These findings highlight the potential of immersive technologies to raise public awareness of conditions like AMD and pave the way for fine-grained, multimodal studies of user behavior to design more impactful experiences.

This work was presented at the IMX conference [20] and in a workshop organized by Fedrha [46]. The experience is available under an open CeCILL licence (IDDN.FR.001.170016.000.S.P.2024.000.31230) at the project webpage.



Figure 15: General workflow of the VR experiment: (a) the participant wearing a headset is going through (b) four main scenarios designed to show "what it is like" to live with AMD and interact with others. (c) Multiple types of data on the user experience are collected including skin conductance response, eye tracking, and questionnaires.

# 9 Bilateral contracts and grants with industry

# 9.1 Bilateral contracts with industry

 

 Participants:
 \* Pierre Kornprobst, \* Sebastián Gallardo, Dorian Mazauric (UniCA, Inria, ABS Team).

 Cifre contract with the company Demain Un Autre Jour (directed by Bruno Génuit), in the scope of the PhD of Sebastián Gallardo (co-supervised by Pierre Kornprobst and Dorian Mazauri, June 2023 – May 2026)

# 10 Partnerships and cooperations

## 10.1 International initiatives

10.1.1 Associate Teams in the framework of an Inria International Lab or in the framework of an Inria International Program

#### FUSION

Title: Functional structure of the retina: A physiological and computational approach

Duration: 2024 ->

Coordinator: Adrian Palacios (adrian.palacios@uv.cl)

#### **Partners:**

• Universidad de Valparaiso (Chile)

#### Inria contact: Bruno Cessac

**Summary:** This project aims to gain a better understanding of the retinal response to complex visual stimuli, and to discover the role played by the network of lateral interneurons (amacrine cells) in this response. To achieve this, we will adopt a dual methodology: experimental and computational. Using real-time control and a feedback loop, we will exploit a computer model of the retina to adapt, in real time, visual stimuli to the recorded responses of retinal cells. The experiments will be carried out in Valparaiso (Chile). The Biovision team will develop the control software and extrapolate the results to Inria Sophia-Antipolis, drawing on its recent theoretical advances in retinal modeling. This is a transdisciplinary, international project at the interface between biology, computer science and mathematical neuroscience. Beyond a better understanding of the role of its network structure in the retina's response to complex spatio-temporal stimuli, this project could potentially have an impact on diagnostic methods for neurodegenerative diseases such as Alzheimer's.

## 10.1.2 Participation in other International Programs

#### **ESTHETICS**

Participants: Bruno Cessac, Rodrigo Cofré, Simone Ebert, Francisco Miqueles, Adrian Palacios, Erwan Petit, Jorge Portal-Diaz, Luc Pronzato, Ludovic Sacchelli.

**Title:** Exploring the functional Structure of THe rETIna with Closed loop Stimulation. A physiological and computational approach.

Partner Institution(s): • Facultad de Ciencias, Universidad de Valparaíso, Chile

- Universidad de Valparaíso, Chile
- I3S, CNRS, Sophia-Antipolis, France

#### Date/Duration: 2 years

Additional info/keywords: This project, funded by the Idex RISE fundings, aims to better understand the retina's response to complex visual stimuli and unravel the role played by the lateral interneurons network (amacrine cells) in this response. For this, we will adopt an experimental methodology using real-time control and feedback loop to adapt, in real-time, visual stimulations to recorded retinal cell responses. The experiments will be done in Valparaiso (Chile). The Biovision team will develop the control software and extrapolate results at Inria Sophia-Antipolis based on its recent theoretical advances in retina modeling. This is, therefore, a transdisciplinary and international project at the interface between biology, computer sciences and mathematical neurosciences. Beyond a better understanding of its network structure's role in shaping the retina's response to complex spatiotemporal stimuli, this project could potentially impact methods for diagnosing neurodegenerative diseases like Alzheimer's.

# 10.2 International research visitors

# 10.2.1 Visits of international scientists

#### Other international visits to the team

### **Francisco Miqueles**

Status (intern (master/eng))

Institution of origin: Universidad Valparaiso

Country: Chile

Dates: 01/02/2024-28/05/2024

Context of the visit: collaboration with A. Palacios team, ESTHETICS project.

Mobility program/type of mobility: internship.

#### Jorge Portal Diaz

Status (PhD)

Institution of origin: Universidad Valparaiso

Country: Chile

Dates: 01/02/2024-28/05/2024

Context of the visit: collaboration with A. Palacios team, ESTHETICS project.

Mobility program/type of mobility: Research stay.

## 10.3 National initiatives

Participants: \* Bruno Cessac, \* Pierre Kornprobst, \* Hui-Yin Wu.

## 10.3.1 ANR

### ShootingStar

Title: Processing of naturalistic motion in early vision

Programme: ANR

Duration: April 2021 - March 2025

Coordinator: Mark WEXLER (CNRS-INCC),

**Partners:** 

- Institut de Neurosciences de la Timone (CNRS and Aix-Marseille Université, France)
- Institut de la Vision (IdV), Paris, France
- Unité de Neurosciences Information et Complexité, Gif sur Yvette, France
- Laboratoire Psychologie de la Perception UMR 8242, Paris

Inria contact: Bruno Cessac

**Summary:** The natural visual environments in which we have evolved have shaped and constrained the neural mechanisms of vision. Rapid progress has been made in recent years in understanding how the retina, thalamus, and visual cortex are specifically adapted to processing natural scenes. Over the past several years it has, in particular, become clear that cortical and retinal responses to dynamic visual stimuli are themselves dynamic. For example, the response in the primary visual cortex to a sudden onset is not a static activation, but rather a propagating wave. Probably the most common motions in the retina are image shifts due to our own eye movements: in free viewing in humans, ocular saccades occur about three times every second, shifting the retinal image at speeds of 100-500 degrees of visual angle per second. How these very fast shifts are suppressed, leading to clear, accurate, and stable representations of the visual scene, is a fundamental unsolved problem in visual neuroscience known as saccadic suppression. The new Agence Nationale de la Recherche (ANR) project "ShootingStar" aims at studying the unexplored neuroscience and psychophysics of the visual perception of fast (over 100 deg/s) motion, and incorporating these results into models of the early visual system.

#### DEVISE

**Title:** From novel rehabilitation protocols to visual aid systems for low vision people through Virtual Reality

#### Programme: ANR

Duration: 2021–2025

Coordinator: Eric Castet (Laboratoire de Psychologie Cognitive, Marseille)

#### **Partners:**

- CNRS/Aix Marseille University AMU, Cognitive Psychology Laboratory
- AMU, Mediterranean Virtual Reality Center

#### Inria contact: Pierre Kornprobst

**Summary:** The ANR DEVISE (Developing Eccentric Viewing in Immersive Simulated Environments) aims to develop in a Virtual Reality headset new functional rehabilitation techniques for visually impaired people. A strong point of these techniques will be the personalization of their parameters according to each patient's pathology, and they will eventually be based on serious games whose practice will increase the sensory-motor capacities that are deficient in these patients.

## **CREATTIVE3D**

Title: Creating attention driven 3D contexts for low vision

Programme: ANR

Duration: 2022–2026

Coordinator: Hui-Yin Wu

#### **Partners:**

- Université Côte d'Azur I3S, LAMHESS, CoBTEK laboratories
- CNRS/Aix Marseille University AMU, Cognitive Psychology Laboratory
- **Summary:** CREATTIVE3D deploys virtual reality (VR) headsets to study navigation behaviors in complex environments under both normal and simulated low-vision conditions. We aim to model multimodal user attention and behavior, and use this understanding for the design of assisted creativity tools and protocols for the creation of personalized 3D-VR content for low vision training and rehabilitation.

#### TRACTIVE

Title: Towards a computational multimodal analysis of film discursive aesthetics

**Programme:** ANR

Duration: 2022-2026

Coordinator: Lucile Sassatelli

#### **Partners:**

- Université Côte d'Azur CNRS I3S
- Université Côte d'Azur, CNRS BCL
- Sorbonne Université, GRIPIC
- Université Toulouse 3, CNRS IRIT
- Université Sorbonne Paris Nord, LabSIC

#### Inria contact: Hui-Yin Wu

**Summary:** TRACTIVE's objective is to characterize and quantify gender representation and women objectification in films and visual media, by designing an AI-driven multimodal (visual and textual) discourse analysis. The project aims to establish a novel framework for the analysis of gender representation in visual media. We integrate AI, linguistics, and media studies in an iterative approach that both pinpoints the multimodal discourse patterns of gender in film, and quantitatively reveals their prevalence. We devise a new interpretative framework for media and gender studies incorporating modern AI capabilities. Our models, published through an online tool, will engage the general public through participative science to raise awareness towards gender-in-media issues from a multi-disciplinary perspective.

# **11** Dissemination

\* Alexandre Bonlarron \* Bruno Cessac \* Johanna Delachambre \* Simone Ebert \* Jerôme Emonet \* Pierre Kornprobst \* Erwan Petit \* Florent Robert \* Hui-Yin Wu

#### 11.1 Promoting scientific activities

# 11.1.1 Scientific events: organisation

## Member of the organizing committees

• H.-Y. Wu co-organized with Cécile Mazon (EPI Flowers) the Journées Scientifique d'Inria on "Handicap, perte d'autonomie & numérique" in Paris, November 19-20

## 11.1.2 Scientific events: selection

#### Member of the conference program committees

• H.-Y. Wu was associate chair for the conference ACM Interactive Media Experiences (IMX 2024).

#### Reviewer

- B. Cessac has been a reviewer for the 46th Annual meeting of the IEEE "Engineering in Medicine and Biological Society".
- P. Kornprobst was a reviewer for The 33nd International Joint Conference on Artificial Intelligence (IJCAI-24), track IJCAI 2024 Main Track.
- H.-Y. Wu was a reviewer for the IEEE Conference on Virtual Reality (IEEEVR 2024).
- H.-Y. Wu was a reviewer for the IEEE International Symposium on the Internet of Sounds
- H.-Y. Wu was a reviewer for the ACM Symposium on Engineering Interactive Computing Systems (EICS)

#### 11.1.3 Journal

#### **Reviewer - reviewing activities**

- H.-Y. Wu was a reviewer for the journal Elsevier Displays.
- H.-Y. Wu was a reviewer for the journal Springer Multimedia Tools and Applications.

#### 11.1.4 Invited talks

- B. Cessac gave a talk in the Insitut de Physique de Nice, entitled "The retina from a physical perspective."
- H.-Y. Wu gave an invited talk at the 48th "Forum de culture scientifique" about her research in virtual reality and artificial intelligence.
- H.-Y. Wu participated in the round table for the Educazur/France Immersive Learning event on "Immersive technologies and learning".

## 11.1.5 Scientific expertise

- P. Kornprobst reviewed one proposition submitted to the ERC Advanced Grant 2023 Call.
- H.-Y. Wu was part of the Maître de Conférences selection committee for UniCA DS4H.

#### 11.1.6 Research administration

- B. Cessac was a member of the Comité Scientifique de l'Institut Neuromod.
- B. Cessac is a member of the Bureau du Comité des Equipes Projets.
- B. Cessac was a member of the Comité de Suivi et Pilotage Spectrum.
- H.-Y. Wu is a member of the Comité de Suivi Doctoral of Inria.

## 11.2 Teaching - Supervision - Juries

#### 11.2.1 Teaching

- Licence 2: A. Bonlarron (58 hours, TD) *Mécanismes internes des systèmes d'exploitation*, Licence en informatique, UniCA, France.
- Master 1: B. Cessac (24 hours, cours) *Introduction to Modelling in Neuroscience*, master Mod4NeuCog, UniCA, France.
- Master 1: E. Petit (15 hours, TD), *Introduction to Modelling in Neuroscience*, master Mod4NeuCog, UniCA, France.

- License 1: E. Petit (36 hours TD), *Bases de l'informatique en python*, License math/info, UniCA, France.
- License 2: J. Emonet (22 hours), Introduction à l'informatique, License SV, UniCA, France.
- License 3: J. Emonet (16 hours), *Programmation python et environnement linux*, License BIM, UniCA, France.
- License 3: J. Emonet (20 hours), Biostatistiques, License SV, UniCA, France.
- Master 1: J. Delachambre (17.5 hours TD), *Création de mondes virtuels*, Master en Informatique (SI4), Polytech Nice Sophia, UniCA, France
- Master 2: J. Delachambre (9 hours TD), *Techniques d'interaction et multimodalité*, Master en Informatique (SI5) mineure IHM, Polytech Nice Sophia, UniCA, France
- Master 1: R. Robert (30 hours TD), *Création de mondes virtuels*, Master en Informatique (SI4), Polytech Nice Sophia, UniCA, France
- Licence 3: F. Robert (39.5 hours TD), Analyse du besoin & Application Web, MIAGE, UniCA, France
- Master 1: H-Y. Wu (5 hours course), *Introduction to scientific research*, Master 1, DS4H, UniCA, France
- Master 1: H-Y. Wu (8 hours course), *Introduction to scientific research*, Master en Informatique (SI5), Polytech Nice Sophia, UniCA, France
- Master 1: H-Y. Wu (12 hours course), *Création de mondes virtuels*, Master en Informatique (SI4), Polytech Nice Sophia, UniCA, France
- Master 2: H-Y. Wu (6 hours course), *Techniques d'interaction et multimodalité*, Master en Informatique (SI5) mineure IHM, Polytech Nice Sophia, UniCA, France
- Master 2: H-Y. Wu supervises M2 student final projects (TER), 2 hours, Master en Informatique (SI5), Polytech Nice Sophia, UniCA, France

#### 11.2.2 Supervision

- B. Cessac co-supervised the PhD of Jérôme Emonet (with A. Destexhe), "A retino-cortical model to study movement-generated waves in the visual system", defended on November 28th.
- B. Cessac supervises the PhD of Erwan Petit, "Modeling activity waves in the retina".
- B. Cessac co-supervised the M1 and M2 internship of Laura Piovano, Master Mod4NeuCog, in collaboration with S. Ebert.
- B. Cessac, P. Kornprobst and H. Wu co-supervised the M1 and M2 intership of C. Kyriazis, Master Mod4NeuCog.
- P. Kornprobst co-supervised (with J.-C. Régin) the PhD of Alexandre Bonlarron, "Pushing the limits of reading performance screening with Artificial Intelligence: Towards large-scale evaluation protocols for the Visually Impaired".
- P. Kornprobst co-supervises (with D. Mazauric) the PhD of Sebastian Gallardo, "Making newspaper layouts dynamic : Study of a new combinatorial/ geometric packing problems". CIFRE contract with the company Demain un Autre Jour (Toulouse).
- P. Kornprobst and H.-Y. Wu co-supervise the PhD of Johanna Delachambre, "Social interactions in low vision: A collaborative approach based on immersive technologies".

- H.-Y. Wu co-supervised (with Marco Winckler and Lucile Sassatelli) the PhD of Florent Robert, "Analyzing and Understanding Embodied Interactions in Extended Reality Systems", defended on December 9, 2024.
- H.-Y. Wu co-supervises (with Lucile Sassatelli) the PhD of Franz Franco Gallo, "Modeling 6DoF Navigation and the Impact of Low Vision in Immersive VR Contexts".
- H.-Y. Wu co-supervises (with Marco Winckler and Aline Menin) the PhD of Clément Quéré, "Apports des réalités étendues pour l'exploration visuelle de données spatio-temporelles".
- H.-Y. Wu co-supervises (with Lucile Sassatelli and Frédéric Précioso) the PhD of Julie Tores, "Deep learning to detect objectification in films and visual media".
- H.-Y. Wu co-supervises (with Marco Winckler) the PhD of Pauline Devictor, "Empathetic storytelling in interactive extended reality".
- H.-Y. Wu supervised the Masters 2 internship of Pauline Devictor.

#### 11.2.3 Juries

- B. Cessac was a member of the comité de thèse of Anastasiia Maslianitsyna (Institut de la Vision, Paris).
- H.-Y. Wu was a jury member for the defense of Paritosh Sharma (University of Paris-Saclay, Paris)
- H.-Y. Wu was a member of the comité de thèse of Géraldine Cherchi (UniCA, SVS, Nice)
- H.-Y. Wu was a member of the comité de thèse of Kevin Galéry (UniCA, SVS, Nice)
- H.-Y. Wu was a member of the comité de thèse of Yujie Huang (Université de Nantes, France)

# 11.3 Popularization

#### 11.3.1 Other science outreach relevant activities

• H.-Y. Wu gave an outreach talk on pursuing a scientific career in computer science to students at Lycée Calmette, Nice

# 12 Scientific production

### 12.1 Major publications

- [1] A. Calabrèse, V. Fournet, S. Dours, F. Matonti, E. Castet and P. Kornprobst. 'A New Vessel-Based Method to Estimate Automatically the Position of the Nonfunctional Fovea on Altered Retinography From Maculopathies'. In: *Translational vision science & technology* 12 (7th July 2023). DOI: 10.116 7/tvst.12.7.9. URL: https://inria.hal.science/hal-04159981.
- [2] E. Castet, J. Termoz-Masson, S. Vizcay, J. Delachambre, V. Myrodia, C. Aguilar, F. Matonti and P. Kornprobst. 'PTVR A software in Python to make virtual reality experiments easier to build and more reproducible'. In: *Journal of Vision* 24 (23rd Apr. 2024). DOI: 10.1167/jov.24.4.19. URL: https://inria.hal.science/hal-04579084.
- [3] B. Cessac, I. Ampuero and R. Cofré. 'Linear response for spiking neuronal networks with unbounded memory'. In: *Entropy* 23.2 (Feb. 2021). L'institution a financé les frais de publication pour que cet article soit en libre accès, p. 155. DOI: 10.3390/e23020155. URL: https://hal.inria.fr/hal-01895095.
- [4] J. Delachambre, H.-Y. Wu, S. Vizcay, M. Di Meo, F. Lagniez, C. Morfin-Bourlat, S. Baillif and P. Kornprobst. 'AMD Journee: A Patient Co-designed VR Experience to Raise Awareness Towards the Impact of AMD on Social Interactions'. In: IMX 2024 ACM International Conference on Interactive Media Experiences. Stockholm, Sweden, June 2024. DOI: 10.1145/3639701.3656314. URL: https://hal.science/hal-04573093.

- [5] S. Ebert, T. Buffet, B. Sermet, O. Marre and B. Cessac. 'Temporal pattern recognition in retinal ganglion cells is mediated by dynamical inhibitory synapses'. In: *Nature Communications* 15.1 (1st July 2024), p. 6118. DOI: 10.1038/s41467-024-50506-7. URL: https://inria.hal.scien ce/hal-03939794.
- [6] O. Faugeras, J. Touboul and B. Cessac. 'A constructive mean field analysis of multi population neural networks with random synaptic weights and stochastic inputs'. In: *Frontiers in Computational Neuroscience* 3.1 (2009). DOI: 10.3389/neuro.10.001.2010. URL: http://arxiv.org/abs/080 8.1113.
- [7] F. F. Gallo, H.-Y. Wu and L. Sassatelli. 'Human Trajectory Forecasting in 3D Environments: Navigating Complexity under Low Vision'. In: *ACM Digital Library*. MMVE 2024 - ACM Multimedia Systems Workshop on IMmersive Mixed and Virtual Environment Systems. MMVE '24: Proceedings of the 16th International Workshop on Immersive Mixed and Virtual Environment Systems. Bari, Italy: ACM, 15th Apr. 2024, pp. 57–63. DOI: 10.1145/3652212.3652223. URL: https://inria.hal.sc ience/hal-04569869 (cit. on pp. 12, 21).
- [8] D. Matzakos-Karvouniari, L. Gil, E. Orendorff, O. Marre, S. Picaud and B. Cessac. 'A biophysical model explains the spontaneous bursting behavior in the developing retina'. In: *Scientific Reports* 9.1 (Dec. 2019), pp. 1–23. DOI: 10.1038/s41598-018-38299-4. URL: https://hal.sorbonne-u niversite.fr/hal-02045700.
- [9] N. V. K. Medathati, H. Neumann, G. S. Masson and P. Kornprobst. 'Bio-Inspired Computer Vision: Towards a Synergistic Approach of Artificial and Biological Vision'. In: *Computer Vision and Image Understanding (CVIU)* (Apr. 2016). DOI: 10.1016/j.cviu.2016.04.009. URL: https://hal.inr ia.fr/hal-01316103.
- [10] J. Naudé, B. Cessac, H. Berry and B. Delord. 'Effects of Cellular Homeostatic Intrinsic Plasticity on Dynamical and Computational Properties of Biological Recurrent Neural Networks'. In: *Journal* of Neuroscience 33.38 (2013), pp. 15032–15043. DOI: 10.1523/JNEUROSCI.0870-13.2013. URL: https://hal.inria.fr/hal-00844218.
- [11] F. A. S. Robert, H.-Y. Wu, L. Sassatelli, S. Ramanoel, A. Gros and M. Winckler. 'An Integrated Framework for Understanding Multimodal Embodied Experiences in Interactive Virtual Reality'. In: 2023 IMX ACM International Conference on Interactive Media Experiences. Nantes, France, 12th June 2023, https://dl.acm.org/conference/imx.DOI: 10.1145/3573381.3596150. URL: https://hal.science/hal-04102737 (cit. on pp. 8, 11, 12).
- [12] F.A. S. Robert, H.-Y. Wu, L. Sassatelli and M. Winckler. 'Task-based methodology to characterise immersive user experience with multivariate data'. In: IEEE VR 2024 - 31st IEEE conference on virtual reality and 3D user interfaces. Orlando (FL), United States, 16th Mar. 2024. URL: https://h al.science/hal-04446066 (cit. on pp. 12, 21).
- [13] S. Souihel and B. Cessac. 'On the potential role of lateral connectivity in retinal anticipation'. In: *Journal of Mathematical Neuroscience* 11 (Jan. 2021). DOI: 10.1186/s13408-020-00101-z. URL: https://hal.inria.fr/hal-02929183.
- [14] H.-Y. Wu, A. Calabrèse and P. Kornprobst. 'Towards Accessible News Reading Design in Virtual Reality for Low Vision'. In: *Multimedia Tools and Applications* (16th May 2021). URL: https://hal.inria.fr/hal-03192461 (cit. on p. 8).
- H.-Y. Wu, F. A. S. Robert, T. Fafet, B. Graulier, B. Passin-Cauneau, L. Sassatelli and M. Winckler.
   'Designing Guided User Tasks in VR Embodied Experiences'. In: *Proceedings of the ACM on Human-Computer Interaction* 6.158 (2022), pp. 1–24. DOI: 10.1145/3532208. URL: https://hal.inria.fr/hal-03635452 (cit. on p. 8).

# 12.2 Publications of the year

# International journals

[16] Best Paper

E. Castet, J. Termoz-Masson, S. Vizcay, J. Delachambre, V. Myrodia, C. Aguilar, F. Matonti and P.

Kornprobst. 'PTVR – A software in Python to make virtual reality experiments easier to build and more reproducible'. In: *Journal of Vision* 24 (23rd Apr. 2024). DOI: 10.1167/jov.24.4.19. URL: https://inria.hal.science/hal-04579084 (cit. on p. 7).

- [17] S. Ebert, T. Buffet, B. Sermet, O. Marre and B. Cessac. 'Temporal pattern recognition in retinal ganglion cells is mediated by dynamical inhibitory synapses'. In: *Nature Communications* 15.1 (1st July 2024), p. 6118. DOI: 10.1038/s41467-024-50506-7. URL: https://inria.hal.scien ce/hal-03939794 (cit. on p. 13).
- [18] E. Kartsaki, G. Hilgen, E. Sernagor and B. Cessac. 'How does the inner retinal network shape the ganglion cells receptive field : a computational study'. In: *Neural Computation* 36.6 (1st June 2024), pp. 1041–1083. DOI: 10.1162/neco\_a\_01663. URL: https://hal.science/hal-04161982 (cit. on p. 15).
- [19] S. Le Briz and H.-Y. Wu. 'Theatron VR AI : (R)amener au théâtre des publics non-initiés ou des pièces sorties des répertoires : VR et arts de l'improvisation'. In: *Loxias*. Cent ans après Kafka 85. (July 2024). URL: https://hal.science/hal-04634335.

#### International peer-reviewed conferences

- [20] J. Delachambre, H.-Y. Wu, S. Vizcay, M. Di Meo, F. Lagniez, C. Morfin-Bourlat, S. Baillif and P. Kornprobst. 'AMD Journee: A Patient Co-designed VR Experience to Raise Awareness Towards the Impact of AMD on Social Interactions'. In: IMX 2024 ACM International Conference on Interactive Media Experiences. Stockholm, Sweden, June 2024. DOI: 10.1145/3639701.3656314. URL: https://hal.science/hal-04573093 (cit. on pp. 24, 25).
- [21] F. F. Gallo. 'DiVR: incorporating context from diverse VR scenes for human trajectory prediction'. In: CV4Metaverse workshop - 3rd Computer Vision for Metaverse Workshop / Co-located at ECCV 2024 - European Conference on Computer Vision. Ed. by H.-Y. Wu and L. Sassatelli. Milano, Italy, 29th Sept. 2024. URL: https://hal.science/hal-04771825 (cit. on p. 21).
- [22] F. F. Gallo, H.-Y. Wu and L. Sassatelli. 'Human Trajectory Forecasting in 3D Environments: Navigating Complexity under Low Vision'. In: ACM Digital Library. MMVE 2024 - ACM Multimedia Systems Workshop on IMmersive Mixed and Virtual Environment Systems. MMVE '24: Proceedings of the 16th International Workshop on Immersive Mixed and Virtual Environment Systems. Bari, Italy: ACM, 15th Apr. 2024, pp. 57–63. DOI: 10.1145/3652212.3652223. URL: https://inria.hal.sc ience/hal-04569869.
- [23] C. Quere, A. Menin, R. Julien, H.-Y. Wu and M. Winckler. 'HandyNotes: using the hands to create semantic representations of contextually aware real-world objects'. In: *IEEE Xplore*. IEEE VR 2024
   The 31st IEEE Conference on Virtual Reality and 3D User Interfaces. Orlando, Florida, United States, 16th Mar. 2024. URL: https://inria.hal.science/hal-04425616 (cit. on p. 24).
- [24] F. A. S. Robert, H.-Y. Wu, L. Sassatelli and M. Winckler. 'Task-based methodology to characterise immersive user experience with multivariate data'. In: IEEE VR 2024 - 31st IEEE conference on virtual reality and 3D user interfaces. Orlando (FL), United States, 16th Mar. 2024. URL: https://h al.science/hal-04446066.
- [25] J. Tores, L. Sassatelli, H.-Y. Wu, C. Bergman, L. Andolfi, V. Ecrement, F. Precioso, T. Devars, M. Guaresi, V. Julliard and S. Lecossais. 'Visual Objectification in Films: Towards a New AI Task for Video Interpretation'. In: CVPR 2024. Vol. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2024, pp. 10864-10874. Seattle (USA, United States, June 2024, pp. 10864–10874. DOI: 10.48550/arxiv.2401.13296. URL: https://hal.science/hal-04840507 (cit. on pp. 12, 22).

#### **Conferences without proceedings**

[26] J. Emonet, A. Destexhe, F. Chavane and B. Cessac. 'A retino-cortical model of anticipation'. In: Neuromod meeting 2024. Antibes, France, 3rd July 2024. URL: https://hal.science/hal-0475 3888 (cit. on p. 15).

#### Doctoral dissertations and habilitation theses

[27] J. Emonet. 'A retino-cortical model to study movement-generated waves in the visual system'. Université côte d'azur, 28th Nov. 2024. URL: https://hal.science/tel-04827484.

#### **Reports & preprints**

- [28] J. Emonet, S. Souihel, M. Di Volo, A. Destexhe, F. Chavane and B. Cessac. A chimera model for motion anticipation in the retina and the primary visual cortex. 26th Sept. 2024. URL: https://in ria.hal.science/hal-04709925.
- [29] S. Gallardo, B. Génuit, D. Mazauric and P. Kornprobst. A Clustering Based Article Template Recommendation System for Newspaper Editors. RR-9560. Universite cote d'Azur, Nov. 2024, p. 32. URL: https://inria.hal.science/hal-04801949 (cit. on p. 19).
- [30] C. Kyriazis, B. Cessac, H.-Y. Wu and P. Kornprobst. Optical Flow Estimation pre-training with simulated stage II retinal waves. RR-9562. Inria & Université Cote d'Azur, CNRS, I3S, Sophia Antipolis, France, Nov. 2024. URL: https://hal.science/hal-04808775 (cit. on p. 15).

#### Other scientific publications

- [31] A. Calabrese, A. Bonlarron, J.-C. Régin and P. Kornprobst. 'A new method to generate automatically highly standardized reading chart text in multiple languages: the case of MNREAD'. In: ARVO 2024 Annual Meeting of the Association for Research in Vision and Ophthalmology. Seattle (USA), Washington, United States, 5th May 2024. URL: https://hal.science/hal-04576614 (cit. on p. 18).
- [32] J. Emonet, A. Destexhe, F. Chavane and B. Cessac. 'A retino-cortical model of anticipation'. In: ICMNS 2024 - International Conference on Mathematical Neuroscience. Dublin, Ireland, 11th June 2024. URL: https://hal.science/hal-04753759 (cit. on pp. 15, 17).
- [33] J. Emonet, A. Destexhe, F. Chavane and B. Cessac. 'A retino-cortical model of anticipation'. In: AREADNE 2024 - Research in Encoding And Decoding of Neural Ensembles conference. Milos, Greece, 25th June 2024. URL: https://hal.science/hal-04705784 (cit. on p. 15).
- [34] C. Kyriazis. 'Can Simulated Retinal Wave pre-training improve Optical Flow Estimation?' In: Neuromod meeting 2024. Ed. by B. Cessac, H.-Y. Wu and P. Kornprobst. Antibes, France, 3rd July 2024. URL: https://inria.hal.science/hal-04735162.
- [35] L. Piovano, S. Ebert and B. Cessac. 'Can Linear-Nonlinear Models capture simple forms of retinal adaptation?' In: NeuroMod 2024 Annual Meeting of the Neuromod Institute. Antibes, France, 3rd July 2024. URL: https://inria.hal.science/hal-04708984.

#### Scientific popularization

[36] F. F. Gallo, H.-Y. Wu and L. Sassatelli. 'Cross-Modal Knowledge Distillation for Human Trajectory Prediction in Virtual Reality'. In: CV4Metaverse 2024 - 3rd Computer Vision for Metaverse Workshop / Co-located at ECCV 2024 - European Conf. on Computer Vision. Milano, Italy, 29th Sept. 2024. URL: https://hal.science/hal-04771856 (cit. on p. 21).

# 12.3 Cited publications

- [37] G. Benvenuti, S. Chemla, A. Boonman, L. Perrinet, G. S. Masson and F. Chavane. 'Anticipatory responses along motion trajectories in awake monkey area V1'. In: *bioRxiv* (2020). DOI: 10.1101/2 020.03.26.010017. URL: https://www.biorxiv.org/content/early/2020/03/29/2020.0 3.26.010017 (cit. on p. 15).
- [38] M. Berry, I. Brivanlou, T. Jordan and M. Meister. 'Anticipation of moving stimuli by the retina'. In: *Nature* 398.6725 (1999), pp. 334–338 (cit. on p. 15).
- [39] J. Besharse and D. Bok. *The Retina and its Disorders*. Elsevier Science, 2011 (cit. on p. 3).

- [40] A. Bhowmick and S. M. Hazarika. 'An insight into assistive technology for the visually impaired and blind people: state-of- the-art and future trends'. In: *J Multimodal User Interfaces* 11 (2017), pp. 149–172. DOI: 10.1007/s12193-016-0235-6 (cit. on p. 4).
- [41] A. Bonlarron, A. Calabrese, P. Kornprobst and J.-C. Régin. 'Constraints First: A New MDD-based Model to Generate Sentences Under Constraints'. In: *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence*. Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence. Macao SAR, China, Aug. 2023, pp. 1893–1901. DOI: 10.24963 /ijcai.2023/210ïżj. URL: https://inria.hal.science/hal-04216267 (cit. on p. 18).
- [42] J. C. Brown, J. E. Goldstein, T. L. Chan, M. R., P. Ramulu and L. V. R. N. S. Group. 'Characterizing functional complaints in patients seeking outpatient low-vision services in the United States'. In: *Ophthalmology* 121.8 (2014), pp. 1655–62. DOI: 10.1016/j.ophtha.2014.02.030 (cit. on p. 4).
- [43] A. Calabrèse, V. Fournet, S. Dours, F. Matonti, E. Castet and P. Kornprobst. 'A New Vessel-Based Method to Estimate Automatically the Position of the Nonfunctional Fovea on Altered Retinography From Maculopathies'. In: *Translational vision science & technology* 12 (July 2023). DOI: 10.1167/t vst.12.7.9. URL: https://inria.hal.science/hal-04159981 (cit. on p. 7).
- [44] B. Cessac and D. Matzakou-Karvouniari. 'The non linear dynamics of retinal waves'. In: *Physica D: Nonlinear Phenomena* (Nov. 2022). DOI: 10.1016/j.physd.2022.133436. URL: https://inria.hal.science/hal-03485137 (cit. on p. 8).
- S. T. L. Chung. 'Cortical Reorganization after Long-Term Adaptation to Retinal Lesions in Humans'. In: *J Neurosci.* 33.46 (2013), pp. 18080–18086. DOI: 10.1523/JNEUROSCI.2764-13.2013 (cit. on p. 4).
- [46] J. Delachambre, M. D. Meo, F. Lagniez, C. Morfin-Bourlat, S. Baillif, H.-Y. Wu and P. Kornprobst. 'À travers les yeux d'un patient malvoyant : une expérience de sensibilisation axée sur les interactions sociales en réalité virtuelle'. In: 2023 Journée Scientifique de la Fedrha sur le thème "Handicap visuel et participation sociale : les apports de la recherche". Lyon, France, Oct. 2023. URL: https://inria.hal.science/hal-04206129 (cit. on p. 25).
- [47] J. L. Fontenot, M. D. Bona, M. A. Kaleem, W. M. McLaughlin, A. R. Morse, T. L. Schwartz, J. D. Shepherd and M. L. Jackson. 'Vision Rehabilitation Preferred Practice Pattern'. In: *Ophthalmology* 125.1 (2018), pp. 228–278. DOI: 10.1016/j.ophtha.2017.09.030 (cit. on p. 4).
- [48] S. Gallardo, M. C. Riff, D. Mazauric and P. Kornprobst. 'Newspaper Magnification with Preserved Entry Points'. working paper or preprint. Sept. 2023. URL: https://hal.science/hal-04210840 (cit. on p. 19).
- [49] Q. Guimard, F. Robert, C. Bauce, A. Ducreux, L. Sassatelli, H.-Y. Wu, M. Winckler and A. Gros.
  'On the link between emotion, attention and content in virtual immersive environments'. In: *ICIP 2022 - IEEE International Conference on Image Processing*. Bordeaux, France, Oct. 2022. DOI: 10.1109/ICIP46576.2022.9897903.URL: https://hal.science/hal-03825059 (cit. on p. 8).
- [50] Q. Guimard, F. Robert, C. Bauce, A. Ducreux, L. Sassatelli, H.-Y. Wu, M. Winckler and A. Gros.
   'PEM360: A dataset of 360° videos with continuous Physiological measurements, subjective Emotional ratings and Motion traces'. In: *MMSys 2022 13th ACM Multimedia Systems Conference*. Athlone, Ireland, June 2022. DOI: 10.1145/3524273.3532895. URL: https://hal.science/hal -03710323 (cit. on p. 8).
- [51] G. Hilgen, E. Kartsaki, V. Kartysh, B. Cessac and E. Sernagor. 'A novel approach to the functional classification of retinal ganglion cells'. In: *Open Biology*. Open Biol. 12.210367 (Mar. 2022). DOI: 10.1101/2021.05.09.443323. URL: https://inria.hal.science/hal-03223562 (cit. on p. 13).
- [52] D. Jancke, W. Erlaghen, G. Schöner and H. Dinse. 'Shorter latencies for motion trajectories than for flashes in population responses of primary visual cortex'. In: *Journal of Physiology* 556 (2004), pp. 971–982 (cit. on p. 15).

- [53] E. Kartsaki. 'How specific classes of retinal cells contribute to vision : a computational model'. Theses. Université Côte d'Azur ; Newcastle University (Newcastle upon Tyne, Royaume-Uni), Mar. 2022. URL: https://theses.hal.science/tel-03869570 (cit. on p. 8).
- [54] G. E. Legge and S. T. L. Chung. 'Low Vision and Plasticity: Implications for Rehabilitation'. In: Annu Rev Vis Sci. 2 (2016), pp. 321–343. DOI: 10.1146/annurev-vision-111815-114344 (cit. on p. 4).
- [55] J. S. Mansfield, S. J. Ahn, G. E. Legge and A. Luebker. 'A new reading-acuity chart for normal and low vision'. In: Ophthalmic and Visual Optics/Noninvasive Assessment of the Visual System Technical Digest, (Optical Society of America, Washington, DC., 1993.) 3 (1993), pp. 232–235 (cit. on p. 18).
- [56] G. Rees, C. L. Saw, E. L. Lamoureux and J. E. Keeffe. 'Self-management programs for adults with low vision: needs and challenges'. In: *Patient Educ Couns* 69.1-3 (Dec. 2007), pp. 39–46. DOI: 10.1016/j.pec.2007.06.016 (cit. on p. 4).
- [57] E. H. Ryan, L. A. Lam, C. M. Pulido, S. R. Bennett and A. Calabrèse. 'Reading Speed as an Objective Measure of Improvement Following Vitrectomy for Symptomatic Vitreous Opacities'. In: *Ophthalmic Surg Lasers Imaging Retina* 51.8 (Aug. 2020), pp. 456–466 (cit. on p. 18).
- [58] S. Souihel and B. Cessac. 'On the potential role of lateral connectivity in retinal anticipation'. In: *Journal of Mathematical Neuroscience* 11 (Jan. 2021). DOI: 10.1186/s13408-020-00101-z. URL: https://hal.inria.fr/hal-02929183 (cit. on p. 8).
- [59] Z. Teed and J. Deng. 'RAFT: Recurrent All-Pairs Field Transforms for Optical Flow'. In: ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part II. Glasgow, United Kingdom: Springer-Verlag, pp. 402–419 (cit. on p. 15).
- [60] W. Wong, X. Su, X. Li, C. Cheung, R. Klein, C. Cheng and T. Wong. 'Global prevalence of age-related macular degeneration and diseas, e burden projection for 2020 and 2040: a systematic review and meta-analysis'. In: *Lancet Glob Health* 2.2 (Feb. 2014), pp. 106–116. DOI: 10.1016/S2214-109X(1 3)70145-1 (cit. on p. 4).
- [61] H.-Y. Wu, A. Calabrèse and P. Kornprobst. 'An Open Virtual Reality Toolbox for Accessible News Reading'. In: *ERCIM News*. Assistive Technologies for a More Accessible and Inclusive Society 130 (July 2022). URL: https://inria.hal.science/hal-03764389 (cit. on p. 8).
- [62] H.-Y. Wu, J. Delachambre, L. Sassatelli and M. Winckler. 'Through the Eyes of Women in Engineering'. In: *Texts of Discomfort*. Carnegie Mellon University: ETC Press, Nov. 2021, pp. 387–414. DOI: 10.1184/R1/17071172. URL: https://hal.inria.fr/hal-03515116 (cit. on p. 8).
- [63] H.-Y. Wu, L. Nguyen, Y. Tabei and L. Sassatelli. 'Evaluation of deep pose detectors for automatic analysis of film style'. In: 10th Eurographics Workshop on Intelligent Cinematography and Editing. Ed. by R. Ronfard and H.-Y. Wu. Reims, France: Association for Computing Machinery (ACM), Apr. 2022. URL: https://inria.hal.science/hal-03634624 (cit. on p. 8).