Bio-Inspired Visual Sensors for Robotic and Automotive Applications

Stefano Mafrica

AIX-MARSEILLE UNIVERSITÉ

FACULTÉ DES SCIENCES DU SPORT DE LUMINY

ED 463 - École Doctorale des Sciences du Mouvement Humain

UMR 7287 - Institut des Sciences du Mouvement

163 Avenue de Luminy, 13009 Marseille, France

Thèse de Doctorat

PhD Thesis

présentée pour obtenir le grade universitaire de docteur en

presented to obtain the grade of philosophical doctor in

Robotique Bio-inspirée

Bio-inspired Robotics

Stefano MAFRICA

Capteurs Visuels Bio-Inspirés pour des Applications Robotiques et Automobiles

Bio-Inspired Visual Sensors for Robotic and Automotive Applications

Soutenue le 12/07/2016 devant le jury :

D. SCARAMUZZA	Prof. Univ. of Zurich (Robotics & Perception Lab)	Rapporteur
P. BONNIFAIT	Prof. Univ. Tech. Compiegne (HEUDIASYC Lab)	Rapporteur
B. MIRAMOND	Prof. Univ. Nice Sophia Antipolis (LEAT Lab)	Examinateur
M. KOVAC	DR Imperial College of London (Aerial Robotics Lab)	Examinateur
C. POSCH	Prof. Univ. Pierre et Marie Curie (Institut de la Vision)	Examinateur
C. BOURDIN	Prof. Univ. Aix-Marseille (ISM Lab)	Examinateur
F. RUFFIER	CR-HDR CNRS (ISM Lab)	Directeur de thèse
A. SERVEL	ADAS & ITS Senior Expert PSA Groupe	Tuteur industriel

Simplicity is the extreme degree of sophistication

Leonardo da Vinci

All art is but imitation of nature

Lucius Annaeus Seneca

Acknowledgments

I would like, first of all, to thank my scientific and industrial supervisors, Franck Ruffier and Alain Servel, for the great opportunity they gave me to learn and work in the fields of biorobotics and automobiles, coupling my passion for robotics and automobiles with the research in innovative ideas coming from nature. I thank them for their constant support, multiple suggestions and high availability which allowed me to carry out and present this work successfully.

I would like to express my gratitude to the members of the committee for accepting to evaluate this thesis, and to all the colleagues of the Biorobotics team at ISM in Marseille and the colleagues of the ADAS, SCFH and SIEP teams at PSA in Velizy for explicitly or implicitly helping me in this great experience.

I especially want to thank Marc Boyron and Julien Dipéri, who have greatly contributed to all the achievements of this thesis, thanks to their technical assistance and their daily availability, kindness and helpfulness. I also thank Stéphane Viollet, Thibaut Raharijaona and Michael Wiertlewski for all the interesting discussions we had together, as well as for their scientific advices, and Julien Serres for allowing me to teach with him a Matlab class at the École des Mines de Saint Etienne.

I sincerely want to thank Jocelyn Monnoyer for his invaluable presence along this experience, giving me both professional and human support, as well as Mihai Socoliuc and Anca Stratulat for their great assistance and helpfulness. I also thank Stéphanie Godiot and Mohsine Menouni for their expertise and help during the work we carried out together.

Hoping to not forget anyone, I would like to thank all the PhD and MSc students, friends of mine, with whom I have worked and rubbed shoulders during these three years, living both serious and amusing moments. Among them, I think about Roman Goulard, Eric Venhoutte, Fabien Colonnier, Augustin Manecy, Fabien Expert, Bruno Boisseau, Aissa Yagoub, Giacomo Rossitto, Gianluigi Brogna, but there are for sure many others that I am missing who should be mentioned.

Last but not the least, I thank with all my heart the people who have been close to me during these three years, without whom this work would not have existed, in particular, my lovely Ethiopian sunshine for her support and patience, and my grandmother who recently passed away and to whom I dedicate this work.

Abstract

Thanks to the advances in the fields of robotics and intelligent transportation systems (ITS), the autonomous vehicles of the future are gradually becoming a reality. As autonomous vehicles will have to behave safely in presence of other vehicles, pedestrians and other fixed and moving objects, one of the most important things they need to do is to effectively perceive both their motion and the environment around them.

In this thesis, we first investigated how bio-inspired visual sensors, called Local Motion Sensors (LMSs), giving 1-D optic flow using a few pixels based on the findings on the fly's visual system, could be used to improve automatic parking maneuvers. For this purpose, we developed a low computational-cost method for detecting and tracking a parking spot in real time using only 1-D OF measurements around the vehicle together with the vehicle's longitudinal velocity and steering angle. Highly simplified 2-D parking simulations were first performed using Matlab/Simulink software, then some preliminary experiments were carried out using a vehicle equipped with two 6-pixel LMSs.

As the main challenge for visual sensors is to correctly operate in high-dynamic-range lighting conditions, we also dealt here with a novel bio-inspired auto-adaptive silicon retina designed and developed by our laboratory in collaboration with the Center of Particle Physics of Marseille (CPPM). We successfully tested this silicon retina, showing that the novel pixel, called M^2APix, which stands for Michaelis-Menten Auto-Adaptive Pixel, can auto-adapt in a 7-decade range and respond appropriately to step changes up to ± 3 decades, while keeping sensitivity to contrasts as low as 2%.

We subsequently developed and tested a novel optic flow sensor based on this auto-adaptive retina and a new robust method for computing the optic flow, which provides several advantages to previously developed optic flow sensors such as its robustness to light levels, textures and vibrations that can be found while operating on the road. To test the performances of this novel sensor and show how it can be used for robotic and automotive applications such as visual odometry, we constructed a car-like robot, called BioCarBot, which estimates its velocity and steering angle by means of an extended Kalman filter (EKF) using only the optic flow measurements delivered by two downward-facing sensors of this kind. Indoor and outdoor experiments were successfully carried out in a 7-decade light level range and using various textures, showing promising perspectives of these sensors for odometry-based applications.

Publications

Journal Papers

S. Mafrica, A. Servel and F. Ruffier, "Minimalistic optic flow sensors applied to indoor and outdoor visual odometry on a car-like robot," *Bioinspiration & Biomimetics* (in revision)

S. Mafrica, S. Godiot, M. Menouni, M. Boyron, F. Expert, N. Marchand, F. Ruffier and S. Viollet, "A bio-inspired analog silicon retina with Michaelis-Menten auto-adaptive pixels sensitive to small and large changes in light," *Optic Express*, vol. 23, pp. 5614-5635, 2015

Patents

S. Mafrica, F. Ruffier and A. Servel, "Dispositif d'aide à la conduite d'un véhicule, à capteurs de flux optique," *French Patent* (pending N^o 1653714)

S. Mafrica, F. Ruffier, A. Servel and J. Monnoyer, "Dispositif de détection à lentille intégrée dans un élément d'un véhicule, pour la mesure de flux optique," *French Patent* (pending N^o 1650538)

Peer-Reviewed Proceedings

S. Mafrica, A. Servel and F. Ruffier, "Optic-Flow Based Car-Like Robot Operating in a 5-Decade Light Level Range," *IEEE International Conference on Robotics and Automation (ICRA)*, Stockholm, Sweden, 2016

S. Mafrica, A. Servel and F. Ruffier, "Towards an Automatic Parking System Using Bio-Inspired 1-D Optical Flow Sensors," *IEEE International Conference on Vehicular Electronics and Safety (ICVES)*, Yokohama, Japan, 2015

Workshops and Posters

A. Manecy, E. Vanhoutte, S. Mafrica, F. Colonnier, F. Ruffier, N. Marchand, S. Viollet, “X4-MaG and RT-MaG: a low-cost open-source micro-quadrotor based on Real-Time Linux and a new Matlab/Simulink toolbox,” *IROS Aerial Open Source Robotics Workshop*, 2015

S. Mafrica, A. Servel and F. Ruffier, “Bio-Inspired Visual Sensors for the Automobile,” *Journées Nationales de la Recherche en Robotique (JNRR)*, Cap Hornu, France, 2015

S. Mafrica, A. Servel and F. Ruffier, “Towards an Automatic Parking System Using 1-D Optical Flow,” *Innovatives Voiture du Futur*, CNRS, Paris, France, 2015

Contents

Chapter 1

General Introduction

1.1 Context and motivation

Intelligent and autonomous systems are already part of our way of living, whether we consciously perceive it or not. Indeed, thanks to the advances in the fields of robotics and intelligent transportation systems (ITS), the autonomous vehicles of the future are gradually becoming a reality.

As autonomous vehicles will have to behave safely in presence of other vehicles, pedestrians and other fixed and moving objects, one of the most important things they need to do is to effectively perceive both their motion and the environment around them. Sensors are, in fact, crucial for safe and autonomous operation of such systems in unknown and dynamic environments, where it is impossible to have complete "a priori" information. Thus, the number and types of sensing technologies embedded on cars are widely increasing to deliver redundant and complementary information, and eventually achieve the level of autonomy and safety which will be required, for instance, by ISO and IEC standards.

Advanced driver assistance systems (ADAS) available on the market already make use of a number of sensors to obtain as much information as possible about the environment around the vehicle, but they are still far from allowing cars to perform truly autonomous tasks.

Some prototypes of the first autonomous vehicles have already been presented, but such solutions are still too complex and expensive, and often require too large computational resources to be onboard commercial cars. This is the reason why car manufacturers are constantly looking for alternative, effective low-cost sensors that may integrate or replace existing technologies.

Effective technological solutions could be found by looking at the nature, which sometimes suggests a different approach, aiming to process only the information needed for well-defined tasks. The insects, for instance, are able to perform complex and fast maneuvers, and avoid obstacles by using only "low-level" visual information and little computational resources, and they are only an example among thousands.

From these observations, the Biorobotics Team of the Institute of Movement Sciences at the Aix-Marseille University, France has been designing and developing for more than 20 years visual sensors inspired by the insects' vision to measure locally the optical flow, i.e. the velocity of visual patterns on

the retina, with a very few pixels and computational resources.

In the following sections, we will provide an insight of the various fields concerned in this thesis. An introduction on existing intelligent and autonomous systems, together with the limits of the technologies used, will be first provided. Then, some important aspects of human and animal vision will be presented, with a particular attention on the insects' visual system and the optic flow. Lastly, some bio-inspired technologies will be presented to show how biological findings can widely help in the development of new efficient and effective technological solutions.

1.2 Intelligent and autonomous robots

Since Isaac Asimov coined the term "robotics" in its "three laws of robotics" (1942) [1], and Norbert Wiener formulated the principles of "cybernetics" (1948) [2], defining the basis of practical robotics, the concept of intelligent and autonomous machines rapidly evolved up to our days.

One of the first examples of a general-purpose intelligent mobile robot, called Shakey, was created by the Artificial Intelligence Center at the Stanford Research Institute (SRI International) between 1966 and 1972 [3] (Fig. 1.1(a)). While other robots existing at that time would have to be instructed on each step for completing a task, Shakey was able to (i) process the visual scene by means of a television camera and (ii) reason about its own actions by analyzing the task asked and breaking it down into basic commands. The project combined the newest research results in robotics, computer vision and natural language processing, melding for the first time sensing, reasoning and acting. Some of the most remarkable results of the project include the Hough transform, the A^* search algorithm and the visibility graph method, which have been subsequently used in a wide range of applications other than robotic. Indeed, the Shakey robot is an example among thousands showing how research works in robotics can boost scientific studies and industrial developments of new technologies used in various application fields.

Since then, the cost and the size of sensors, actuators and computational units have reduced exponentially, making it possible to develop cheaper and more compact robots with increasing intelligence and autonomy. As as result, the number of application fields of robotic systems widely increased during the last few decades, ranging from automation and manufacturing (the initial main market), to defense and civil surveillance and intervention, to health care and human assistance, to home service and entertainment, to intelligent transportation systems (ITSs). A remarkable evidence of this expansion of robotics could be seen within the last decade, for instance, when Google acquired 7 robotic companies in 2013, like Boston Dynamics and Meka Robotics, and developed the first self-driving car called Google Car.

The main functions characterizing an intelligent and autonomous robot can be defined as follows:

- perception (sensing and processing);
- reasoning (interpretation and decision making);
- control (action and interaction).

Figure 1.1: (a) The Shakey robot created by the Artificial Intelligence Center at the Stanford Research Institute. Adapted from [3]. (b) The Robox cobot (collaborative robot) developed by BlueBotics SA and EPFL Autonomous Systems Lab and shown at the Swiss National Exhibition Expo 2002. Adapted from [4]. (c) The Asimo robot created by Honda with its ancestors.

Effective perception and control represent the basic and necessary capabilities that bring mobile robotic systems closer to autonomy, and may be sufficient for a robot to autonomously operate in relatively simple environments. However, when the tasks and the environments become dynamic and complex, these capabilities have often to be fused with higher level decision-making mechanisms, which focus on more abstract cognitive planning abilities, to bring forth truly autonomous and intelligent systems.

As the number of autonomous agents increases, and the tasks and the environments become more complex, the level of autonomy and intelligence have to be distributed over a swarm of robots in order to cooperate with each other and with humans. Thereby, a new generation of robots, called "Cobots" [4, 5], have been designed and developed to cooperate in close interaction with humans in both public and private spaces (industrial sites, cities, museums, hospitals, etc). Large scale experiments have been successfully carried out in human populated environments using, for instance, tour-guide robots (Fig. 1.1(b)), humanoid robots (Fig. 1.1(c)) and cyber-cars (Fig. 1.6).

This new paradigm represents a major challenge for both industry and human society, and requires new robots' characteristics for balancing safety, efficiency and autonomy constraints, as well as being socially acceptable and intuitive for human-robot interaction and cooperation. This requires some technological breakthroughs to provide autonomous but shared control, while dealing with incompleteness and uncertainty on the environment, increasing robustness and safety, and satisfying real-time and cost constraints using miniaturized embedded systems.

Indeed, if we look at nature we can see that the animals, and particularly the insects, can perform complex tasks such as navigation and hazardous maneuvers with a tiny brain, and therefore with very few processing and reasoning capabilities (see sections 1.6.3, 1.6.4 and 1.7.3). This is possible by using very specialized, effective sensors and actuators, whereas the connection between these, which usually defines the "intelligence" of an agent, can be quite simple depending on the animal's morphology: in this case, we speak about "embodied intelligence".

As in this thesis we mainly deal with the perception, the sensors used in robotics and automobiles, together with their advantages and drawbacks, will be described in details in the following section.

1.3 Sensing technologies in robotics and automobiles

There exist a wide variety of sensors used in mobile robotic systems, which can be divided in 2 main groups: proprioceptive and exteroceptive sensors. Proprioceptive sensors measure values internal to the system, such as wheel/motor sensors (e.g. encoders) and inertial measurements units (IMUs), whereas exteroceptive sensors acquire information external to the system (i.e. from the environment), such as passive and active visual sensors (e.g. cameras or lidars), ultrasonic sensors (sonars), radio-based sensors (radars) and force/tactile sensors (Fig. 1.2).

The sensors can also be divided in active and passive sensors, depending on whether they measure the environmental energy entering the sensor or the environmental reaction after emitting energy into

Figure 1.2: Example of sensors used in robotics and automobiles. (a) Ultrasonic sensor provided by Arduino. (b) Radar sensor provided by Banner Engineering. (c) Three types of lidars provided by Velodyne. (d) Camera provided by Raspberry Pi. (e) Stereo-vision sensor provided by Voltrium Systems. (f) Kinect camera (infrared depth camera) provided by Microsoft.

the environment. Active sensors can often achieve superior performance as they control they the energy emitted into the environment, however they have higher energy consumption and may suffer from interference between its signals and those beyond its control (emitted, for instance, by other nearby robots). Therefore, passive sensors are usually preferred when achieving similar performances to active sensors.

Wheel/motor sensors The wheel/motor sensors measure the internal state (position, velocity, torque, etc) of robot's rotating components (motors or wheels). Probably the most popular sensors of this type are the optical encoders, which measure the angular speed and position of a motor shaft or a wheel. Because these sensors measure the position or velocity of some internal parts of the robot (e.g. wheels), when used for the robot's localization (odometry), they can produce drifts and even completely wrong estimations, for instance, in slippery conditions.

Inertial measurements units (IMUs) The IMUs are electronic devices that measure the linear acceleration, the angular velocity and the orientation of a robot, using a combination of accelerometers, gyroscopes and magnetometers. These sensors are widely used in robotics since they can be very cheap and compact, and provide very good performances for flight stabilization of aerial robots. However, similarly to wheel/motor sensors, when these sensors are used for the robot's localization (odometry), they may suffer from drifts and high-level noise due to the integration over time of acceleration data, especially in presence of strong vibrations.

Global navigation satellite systems (GNSSs) The global navigation satellite systems (GNSSs), such as global positioning systems (GPSs), are space-based external sensors that provide location and time information in all weather conditions, anywhere on the earth where there is an unobstructed line of sight to at least four satellites. They are the only sensors that can deliver directly the robot's absolute position, however, they can operate only in open areas and often suffer from unreliability (wrong or no measurements) and low resolution (up to a few meters). Differential GPSs (DGPSs) can deliver position measurements with much higher resolution (up to a few centimeters) but are still very expensive for use in most of robotic and automotive applications.

Sonars The ultrasonic sensors (or sonars) (Fig. 1.2(a)) are active sensors that measure the time-of-flight (ToF) after the reflection of emitted acoustic waves against an object's surface. They are usually compact and low-cost sensors and have the advantages of detecting almost any type of material. However, they can operate only in a short distance range ($< 5\,m$) and are quite sensitive to the noise that can be produced by the wind or the objects' shape (e.g. thin rod, sharp edge, etc).

Radars The radars (RAdio Detection And Ranging) (Fig. 1.2(b)) are active sensors that use electromagnetic (radio) waves to measure the distance of an object based on the ToF, as well as its velocity based on the Doppler effect (change of frequency of waves emitted from a moving source). They have advantages such as operating in a long range (up to $200\,m$) and providing both position and velocity measurements using 2 different techniques on the same signal. On the other hand, their main disadvantages are the bad angular precision, bad precision for tangential movements and bad reflection on certain surfaces.

Lidars The lidars (LIght Detection and Ranging) (Fig. 1.2(c)), also called laser scanners, are active sensors that use the light (generally infrared) reflected on a rotating mirror to measure the distance of an object based on the ToF. There exist 1-D, 2-D and 3-D scanners, depending on the degree-of-freedom (DOF) of the rotating mirror. Their main advantages are their wide range and high precision, which is also nearly independent of the distance. However, they are very expensive, they comprise some mechanical parts, which make the sensors wear out more easily, and they are sensitive to rain and dust.

Cameras Standard cameras (Fig. 1.2(d)) are widely used in robotics since they are cheap passive sensors that can provide very rich information about the surrounding scene, but they need large computational resources to process the images to extract information about distance and velocity of the objects. In addition, as the 3-D scene is projected onto a 2-D image, cameras require some additional information (from other sensors or from some hypothesis on the environment) to obtain 3-D metric measurements.

Stereo-vision systems (Fig. 1.2(e)) are composed of multiple cameras (usually 2) to retrieve 3-D information on the environment from triangulation, but they are more expensive (multiple cameras looking at the same scene) and need fine calibration and image processing to deliver robust measurements.

Infrared depth cameras, such as the Kinect sensor (Fig. 1.2(f)), give directly range 3-D data using only one infrared camera and a light emitting source therefore can be relatively cheap, but distance information can be provided only in a short range (up to a few meters), and are very sensitive to the lighting conditions and therefore not very suitable for outdoor applications.

1.4 Advanced driver assistance systems

An Advanced Driver Assistance System (ADAS) is a system embedded on a vehicle which is designed to assist the driver by enhancing, adapting and sometimes automating the driving to increase his safety and comfort. Any system easing and facilitating the task of the driver can be therefore considered as an ADAS.

The main tasks of an ADAS system can be listed as follows:

- avoid a dangerous situation that could lead to an accident;
- free the driver of various tasks that could lessen his vigilance;
- assist the driver within his perception of the environment.

The safety features designed to avoid collisions and accidents can be divided in two groups: the technologies that alert the driver to potential problems (passive safety), and those that avoid collisions by implementing safeguards and taking over control of the vehicle (active safety).

ADAS systems that can be found nowadays on commercial cars make use of the sensing technologies presented in section 1.3, sometimes fusing them to achieve higher robustness and accomplish more complex tasks (Fig. 1.3, 1.4).

Since ADAS systems rely on electronics and often include firmware elements, the development of these cutting edge technologies is ruled by international safety standards, such as the IEC-61508 and ISO-26262, which are proposed by specialized technical committees (e.g. ISO/TC 22 for “road vehicles” and ISO/TC 204 for “intelligent transport systems”).

In the following, an insight of the main existing ADAS systems will be given. Some of these technologies have been around for a long time and they have already proven to result in improved driving experiences and better overall road safety. However, a lot of ADAS are right on the cutting edge of emerging automotive technologies as they need an increasing number and types of sensors delivering redundant and complementary information to eventually achieve a higher level of autonomy and safety.

Anti-lock Braking System The anti-lock braking system (ABS) is the first ADAS that has been embedded in commercial cars (by Mercedes in 1978), significantly improving the driving safety (estimated crash reduction from 5% to 85% depending on the road conditions between 1995 and 2007 [6]). The ABS allows the vehicle’s wheels to maintain traction contact with the road surface according to driver’s braking inputs, preventing the wheels from locking up and avoiding uncontrolled skidding, especially on

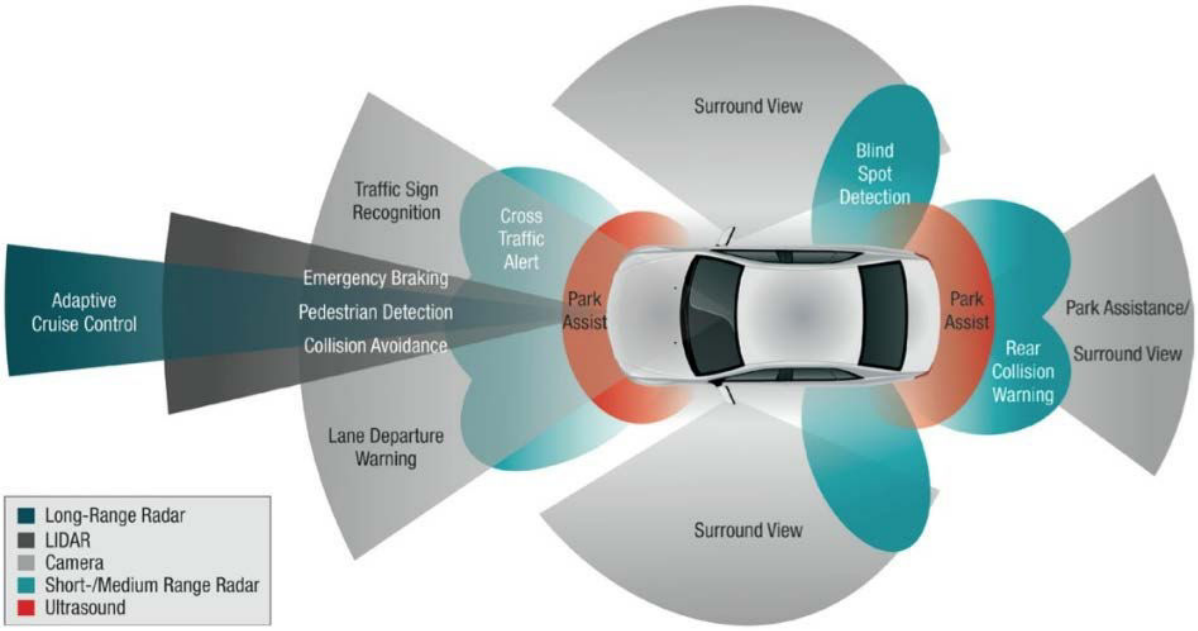

Figure 1.3: Scheme of the sensors used in advanced driver assistance systems (ADAS) available on the market provided by Texas Instrument.

	Sonars	Radars	Lidars	Cameras
Intelligent Speed Adaptation (ISA)				1
Advanced Emergency Brake (AEB)		1-2	1	1
Adaptive Cruise Control (ACC)		1	1	1
Traffic Jam Assist		1	1	1
ACC Stop & Go	4	1-3	1	1
Lane Keeping Assist (LKA)				1
Lane Changing Assist (LCA)		2		2
Blind Spot Detection	4	2		2
Intelligent Parking Assist (IPA)	8	4	1	2
Highway Automated Driving (HAD)		5	2	2

Figure 1.4: Table of the average number and type of sensors used in ADAS systems provided by most of the car manufacturers. Courtesy of Alain Servel.

slippery terrains. To do that, the wheel's angular velocities are compared to the vehicle's velocity estimation derived from other sensors (IMU, GPS, etc.), hence the importance of these sensors' information to be reliable.

Electronic Stability Control The electronic stability control (ESC), also called electronic stability program (ESP), is a sort of extension of the ABS to improve the vehicle's stability by detecting and reducing loss of traction (skidding). The ESC detects loss of steering control (understeering or oversteering) and automatically applies braking commands separately to the 4 wheels to help steer the vehicle where the driver wants to go. Some ESC systems also reduce engine power until control is regained. According to the Insurance Institute for Highway Safety and the U.S. National Highway Traffic Safety Administration (NHTSA), one-third of fatal accidents could be prevented by using this technology [7], which became therefore a mandatory ADAS for any car produced since 2011.

Intelligent Speed Adaptation This ADAS depends on a variety of information to help a driver maintain a legal speed. Since these systems monitor the current speed and compare it with the local speed limit coming from the camera information (traffic signs recognition) and the GPS-cartography information, they can work properly only in certain areas.

Advanced Emergency Braking The advanced emergency braking (AEB) is an ADAS designed to reduce the severity of high speed collisions in the event of a lapse of driver attention. While some recent automatic braking systems can actually prevent collisions, AEB systems are typically meant to slow down the vehicle's velocity to the point where less damage is caused and fatalities are unlikely. Most of the AEB systems that can be found today on cars make use of a radar, a lidar and a camera to detect obstacles and dangerous situations and trigger the emergency braking control.

Adaptive Cruise Control The adaptive cruise control (ACC) is the first ADAS developed to automatically adapt the vehicle's velocity in response to the actions of the vehicles in front of it. The ACC is especially useful on highways where drivers otherwise have to constantly monitor their cruise velocity in function to other vehicles' velocity, and therefore usually automatically shuts down below a certain speed threshold. Newer ADAS systems based on ACC, such as ACC stop & go and the traffic jam assist, have been recently developed by most of the car manufacturers for dense traffic situations and therefore lower velocities and inter-vehicle distances.

Lane Keeping and Changing Assist The lane keeping assist (LKA) system makes sure that the vehicle does not leave its lane accidentally by warning the driver through visual and auditive alerts (lane departure warning), and eventually taking small corrective actions over the driver if this is not reacting. Lane changing assist (LCA) systems go a step further and are actually capable of performing lane changing

maneuvers without any driver action. While LKA systems only need a front-view camera to detect the road lanes, LCA systems additionally need a rear-view camera as well as front and rear radars.

Blind Spot Detection Blind spot detection systems use a variety of sensors to provide a driver with vital information that would be difficult or impossible to come by through any other means. Some of these systems will sound an alarm if they detect the presence of an object within a blind spot, and other include cameras to show images of the blind spot in the head unit monitor.

Intelligent Parking Assist Intelligent parking assist (IPA) systems are designed to help the driver with a parallel and/or perpendicular parking maneuver, but they may vary from one car manufacturer to another. Most of the IPAS systems available on the market involve the use of ultrasonic sensors and a rear-view camera to detect a free parking spot while the driver is driving along a row of parked vehicles [8–10]. Then, some of them simply tell the driver when to turn the steering wheel, to change the gear and to stop the car, while others (most recent) (semi-)automatically control the steering and sometimes the speed based on the estimation of the vehicle's ego-position via odometry. IPAS systems involving Around View Monitor Systems (AVMSs) have been recently proposed [11, 12] to detect and track parking-spot' ground marks during parking maneuvers by means of wide-angle cameras (such as fish-eye and catadioptric cameras) directed toward the ground, giving a view of the surrounding environment.

Highway Automated Driving Systems providing highway automated driving (HAD) experiences combine all the functions of the ADAS previously described with some additional sensors, namely radars and lidars, to make cars autonomously drive on highways and freeways. This is the first and main feature which prototypes of autonomous vehicles have demonstrated (see section 1.5), and therefore it will be probably the first fully-automated driving system included on standard cars within the next few years.

High-performance ADAS systems have been presented in various research works [13–15] but they still require too costly sensors, such as high-resolution lidars and radars as well as large computational resources to reconstruct the surrounding 3-D environment in real time during the maneuvers. Therefore, car manufacturers are constantly looking for alternative low-cost technologies.

First, we notice that, as mentioned at the beginning of this chapter, the more complex are the automated driving features and the higher is the level of uncertainty, the higher is the number and types of sensors used. Second, we notice the importance of vision for improving safety and autonomy of future cars, as all recently-developed ADAS systems make use of at least one camera (Fig. 1.4). As an example, the U.S. NHTSA announced in 2014 that all new vehicles under 4500 kg produced in the U.S. will required to have rear-view cameras by May 2018 [16]. In addtion, night vision systems using both active and passive infrared sensors are gradually integrating ADAS systems to allow drivers to detect objects that are difficult or even impossible to make out at night.

It is also worth noting the importance of an accurate estimation of the vehicle's velocity and pose,

i.e. the odometry, is crucial for the effectiveness of most ADAS systems, from ABS and ESC to IPA and HAD systems. Therefore, recent research works have been presented for performing reliable vehicle's positioning by fusing together the outputs of various onboard systems, such as GPS and ABS, with lane markings detection [17] or 3-D urban maps [18]. Various works based on visual odometry [19, 20] and simultaneous localization and mapping (SLAM) [21, 22] have also been presented, gaining increasing interest in the ITS community.

1.5 Towards intelligent and autonomous vehicles

Although the keyword "autonomous vehicle" has been widely used by most of the car manufactures and by the media only during the last few years, the idea of a self-driving car is definitely not new.

In one of most popular General Motors' exhibitions entitled "Futurama: Highways and Horizons" at the New York's World Fair in 1939, the industrial designer Norman Bel Geddes had the vision of a future featuring automated highways as a solution to traffic congestion. In the solution he presented, the cars were powered by circuits embedded in the roadway and controlled by radio communication. Although Bel Geddes did not call it in this way, that was the first clear vision of the autonomous and connected vehicles of the future.[1]

However, the first driverless car was demonstrated only about 40 years later by the Tsukuba Mechanical Engineering Lab: the car could detect and track white street markers using a black-and-white camera, and reach speeds up to 30 km/h on a dedicated track in Japan [23]. The technology was not ready yet for automatic driving at high speed in highways and freeways, mainly due to the low-frequency image processing and therefore the delays in the driving control feedback.

In 1986, professor Ernst Dickmanns and his research team at the Bundeswehr University of Munich constructed the first complete robotic car, called VaMoRs, based on a Mercedes-Benz van (Fig. 1.5(a)). The 5-ton car was re-engineered such that it was possible to control its steering wheel, throttle, and brakes through electronic commands based on real-time processing of image sequences from and embedded camera. The robotic car was capable of driving in well-marked highways without traffic using saccadic computer vision, probabilistic approaches and parallel computing [24].

One of the greatest challenges in high-speed autonomous driving arises through the rapidly changing visual scenes of the street. Forty years ago, computers were much slower than today, therefore sophisticated computer vision techniques were necessary to obtain useful information in real time. The team of Dickmanns solved the problem through an innovative approach to dynamic vision, called "4-D approach", in which spatio-temporal models were used to estimate 3-D position and velocity components without storing previous images [25]. Inspired by human attention control, the platform carrying the cameras performed saccadic movements to focus the attention on the most relevant details of the visual scene. Kalman filters were extended to perspective imaging and used to achieve robust autonomous driving

[1] After stepping out the exhibition, the visitors proudly wore a pin in which there was written "I have seen the future".

even in presence of noise and uncertainty [25].

Since then, the European Commission and the American Defense Advanced Research Projects Agency (DARPA), among others, started funding major research initiatives with the objective of studying the problem of autonomous navigation in urban and extra-urban environments.

In 1994, as final part of the Eureka PROMETHEUS project, the Dickmanns's team in collaboration with Mercedes-Benz constructed 2 semi-autonomous vision-based cars based on a Mercedes 500 SEL, called VaMP and VITA-2 [26] (Fig. 1.5(b)). The two twin cars could drive more than 1000 km on the highway near the airport Charles-de-Gaulle near Paris, in standard traffic with a speed up to 130 km/h, by automatically keeping a safe distance depending on their speed and passing slower cars in the left lane. The latter required interpreting the road scene also in the rear car's hemisphere, thereby 2 cameras with different focal lengths for each hemisphere (front and rear) have been used in parallel [26].

One year later, a modified version of the VaMP car autonomously drove about 1700 km on public highways from Munich, Germany to Odense, Denmark at a speed up to 180 km/h, automatically passing other cars (lane-changes at a maximum speed of 140 km/h) without GPS and with only 5% of human intervention [26]. This is particularly impressive considering that the system used black-and-white cameras and did not model situations such as road construction sites with yellow lane markings.

Considerable results were subsequently obtained by the research groups of A. Broggi (University of Pavia, Italy) and C. G. Lo Bianco (University of Parma, Italy) between 1996 and 2001 in the context of the ARGO project [27]. A modified version of the Lancia Thema (Fig. 1.5(c)) was able to distinguish traffic lanes, identify ahead vehicles and other interference to its path, without requiring any special road infrastructure, by using only a stereo vision system. The real-world test of the car was done in 1998, when the car completed the "MilleMiglia in Automatico" tour, a journey through Italy along about 2000 km performed in automatic driving with a maximum speed of about 120 km/h, including extra-urban rural roads for the first time ever [27].

In 2004, the first DARPA Grand Challenge, a driverless car competition along a 240-km off-road route in the Mojave Desert region of the United States, took place with the purpose of accelerating the development of autonomous vehicle technologies that could be applied to military requirements. None of the vehicles finished the course, whereas the Sandstorm vehicle (a converted Humvee) of the Carnegie Mellon's Red Team traveled the farthest distance completing only about 12 km of the route [28].

One year later, the second DARPA Grand Challenge involved an off-road route similar to the previous year, but this time 5 vehicles successfully completed the 212-km course, showing the exponential increase of performances of the technologies used. The Stanley vehicle (a modified Volkswagen Touareg) of the Stanford's Racing Team, which won the first place, incorporated measurements from a GPS, a 6DOF IMU, and wheel encoders for the vehicle's pose estimation, whereas the environment was perceived through 4 lidars, a radar system, a stereo-vision and a monocular vision system [29] (Fig. 1.5(d)). The sensors and processing units (7 onboard Pentium M computers) were still too cumbersome and costly to make car manufacturers be interested in such applications.

Figure 1.5: (a) The VaMoRs vehicle based on a Mercedes-Benz van, developed by Dickmanns's research team at Bundeswehr University of Munich in 1986. (b) The VaMP vehicle based on a Mercedes 500 SEL, developed by Dickmanns's research team at Bundeswehr University of Munich in 1994. (c) The ARGO vehicle based on a Lancia Thema, developed by the research groups of A. Broggi (University of Pavia, Italy) and C. G. Lo Bianco (University of Parma, Italy) in 1998. (d) The Stanley vehicle based on a Volkswagen Touareg, developed by the Stanford's Racing Team, the winner of the DARPA Grand Challenge in 2005. (e) The driverless vehicle based on a Piaggio Porter, developed by the Artificial Vision and Intelligent Systems Laboratory (VisLab) at University of Parma, Italy for the Intercontinental Autonomous Challenge (VIAC) in 2010. (f) The Google Car, the first real-build driverless prototype vehicle delivered by Google in 2014.

In 2007, the DARPA Urban Challenge, won by Carnegie Mellon's Tartan Racing Team [14], involved a 96-km urban-area course to be completed in less than 6 hours, while obeying all traffic regulations, merging into traffic and negotiating with other cars and obstacles. Sensor systems become more elegant and complete, and semi-autonomous features begin to hit the mainstream with car manufacturers such as Audi, Volvo, GM and Mercedes incorporating features like collision avoidance, lane keeping and changing, and driver attention assist into their new vehicles.

The VisLab Intercontinental Autonomous Challenge (VIAC) was the challenge conceived by the Artificial Vision and Intelligent Systems Laboratory (VisLab) at University of Parma, Italy, as an extreme test of autonomous vehicles [30]. It ran from July to October 2010, involving 2 driverless Piaggio Porters (Fig. 1.5(e)) driving with virtually no human intervention on an almost 16000 km trip from Parma, Italy to Shanghai, China. The vehicles had to face a plethora of very different and extreme conditions: road, weather, infrastructures, temperature, traffic, and even possibly unlawful behaviors of other vehicles. In addition, the route was unknown since no maps of a large percentage of the trip were available. The first vehicle (leader) was driven autonomously for most of the trip (limited human interventions were needed to define the route and intervene in critical situations) while testing the sensing, decision and control subsystems, and collecting data, whereas the second vehicle (follower) automatically followed the route defined by the leader, requiring no human intervention (100% autonomous) [30]. The sensing technologies used were similar to those used in the DARPA challenges, including 4 lidars, 2 stereo-vision systems, 1 panoramic camera and 1 localization unit, yet being far from integration on standard cars.

The Google Car project lunched in 2009 started testing compact self-driving technologies with the Toyota Prius on freeways in California, moving then with the Lexus RX450h in city streets in 2012. In December 2014, Google delivered the first real-build prototype vehicle which included a GPS, a IMU, wheel encoders, a 3-D lidar on the top, and a radar and a camera at the front (Fig. 1.5(f)). The Google Car have been self-driven up to present over almost 2 million km in the United States, and involved in 12 minor accidents without any people blessed, mostly due to human errors from other drivers.

The Google Car project showed for the first time that autonomous driving was possible both in urban and extra-urban roads, both at low and high speed, using a limited number of on-the-shelf sensors, thanks to their increasing performances and decreasing cost, as well as the increasing embedded intelligence and the vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communication. Therefore, during the last few years several car manufacturers started autonomous vehicle projects, leading to the first prototypes that have been tested on public roads (mainly highways) in Europe, for instance, by the Volkswagen, Mercedes, Volvo and PSA groups. As an example, the vehicle prototype of PSA Peugeot Citroën has been recently driven autonomously along nearly 500 km on highway from Paris to Bordeaux, France.

The Volkswagen's Temporary Auto Pilot [31] received the strongest media attention because it left the impression of a close approach to truly automated driving, although the driver have to continuously watch the forward driving scene to make the system operate. The AutoPilot demonstration proceeded

through the different levels of driver assistance up to the combination of lateral and longitudinal control. The vehicle demonstrated, for instance, the ability to track a slower vehicle in the left adjacent lane and avoid overtaking it on the right, which is illegal in many European countries, as well as to perform an emergency braking maneuver behind a stopped vehicle in its lane.

On parallel lines, the concept of Cybernetic Transport Systems (CTS) or "cybercars", i.e. automated electric vehicles transporting people or merchandises in some protected areas (cities, airports, private tracks, etc), was proposed in 1991 by the french institute INRIA (*Institut National de Recherche en Informatique et en Automatique*). The first successful demonstrations took place at the Schiphol airport of Amsterdam, Netherlands in 1997, with the Cycab vehicles, and in a suburb of Rotterdam, Netherlands in 1998, with the Parkshuttles vehicles provided by 2GetThere (Fig. 1.6(a)). In this case, the features of the autonomous vehicles were completely different than those described above, since they had to drive autonomously at low speed in urban or suburban areas, that is in unknown, unstructured and populated environments.

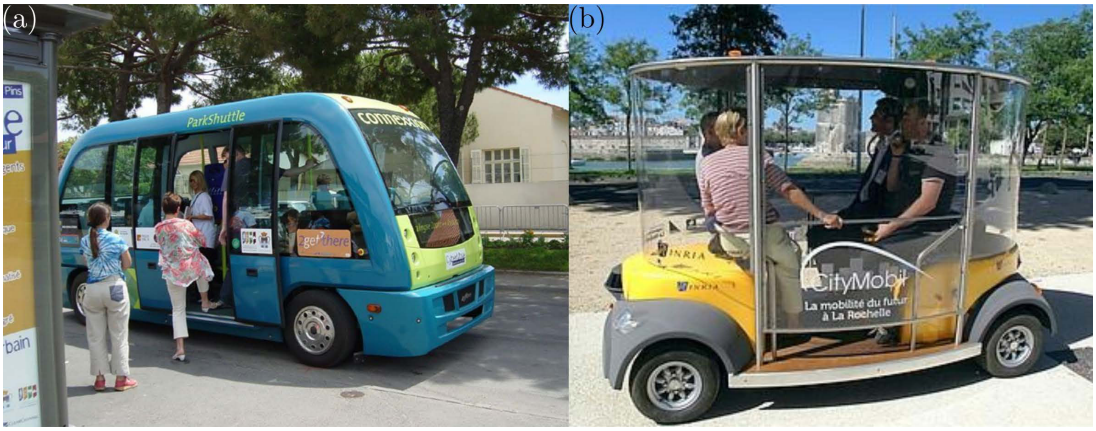

Figure 1.6: (a) The Parkshuttles cybercar developed by INRIA and 2GetThere and demonstrated at the Schiphol airport of Amsterdam, Netherlands in 1997. (b) The Cybus cybercar developed by INRIA and Robosoft and demonstrated in La Rochelle, France in 2011.

Since then, the European Commission started funding major research CTS projects, such as the CyberMobil and CyberMove projects between 2001 and 2011, with the goal of creating a new transportation options to move towards sustainability and increasing the attractiveness of city centers. Today, there exist already some companies that produce cybercars, such as 2GetThere, Robosoft, Induct, Navya and Lohr in Europe, and several early large-scale public experiments have been carried out, such as those with the Cybus vehicles in La Rochelle, France in 2011, developed by INRIA and RoboSoft in the CityMobil EU-funded project (Fig. 1.6(b)).

Several other projects (e.g. Aktiv, SAFESPOT, HAVEit, INTERSAFE 2, SARTRE, InteractIVe) and challenges (e.g. the Multi Autonomous Ground-robotic International Challenge and the Grand Cooperative Driving Challenge) contributed in the scientific and technological advances that are leading most of the car manufacturers to delivering the first commercial (semi-)autonomous cars within a few years.

In 2012, the German Federal Institute for Road Research ("Bundesanstalt für Strabenwesen",BASt) proposed a first formalism of the levels of automated driving divided in five levels [32]. In 2013-2014, the U.S. NHTSA agency, and then the International Society of Automotive Engineers (SAE International) and the International Organization of Motor Vehicle Manufacturers (OICA), have also released their own definition of the levels of automated driving [33–35]. The SAE definition, shown in Fig. 1.7, completely agree with the BASt and OICA definitions, and is therefore commonly accepted and adopted by most of the car manufacturers in Europe since it provides an incremental approach relying on a collaboration between the driver and the system. The existing ADAS systems correspond to the first two levels of automated driving (level 1 and 2 in Fig. 1.7), while prototypes of third and fourth level are already presented by some car manufacturers (level 3, 4 and 5 in Fig. 1.7).

SAE level	Name	Narrative Definition	Execution of Steering and Acceleration/ Deceleration	*Monitoring* of Driving Environment	Fallback Performance of *Dynamic Driving Task*	System Capability *(Driving Modes)*
***Human driver* monitors the driving environment**						
0	**No Automation**	the full-time performance by the *human driver* of all aspects of the *dynamic driving task*, even when enhanced by warning or intervention systems	Human driver	Human driver	Human driver	n/a
1	**Driver Assistance**	the *driving mode*-specific execution by a driver assistance system of either steering or acceleration/deceleration using information about the driving environment and with the expectation that the *human driver* perform all remaining aspects of the *dynamic driving task*	Human driver and system	Human driver	Human driver	Some driving modes
2	**Partial Automation**	the *driving mode*-specific execution by one or more driver assistance systems of both steering and acceleration/ deceleration using information about the driving environment and with the expectation that the *human driver* perform all remaining aspects of the *dynamic driving task*	**System**	Human driver	Human driver	Some driving modes
***Automated driving system* ("system") monitors the driving environment**						
3	**Conditional Automation**	the *driving mode*-specific performance by an *automated driving system* of all aspects of the dynamic driving task with the expectation that the *human driver* will respond appropriately to a *request to intervene*	System	**System**	Human driver	Some driving modes
4	**High Automation**	the *driving mode*-specific performance by an automated driving system of all aspects of the *dynamic driving task*, even if a *human driver* does not respond appropriately to a *request to intervene*	System	System	**System**	Some driving modes
5	**Full Automation**	the full-time performance by an *automated driving system* of all aspects of the *dynamic driving task* under all roadway and environmental conditions that can be managed by a *human driver*	System	System	System	**All driving modes**

Figure 1.7: Table of the levels of automated driving for road vehicles defined by SAE International.

All the presented systems offer a number of degrees of assistance to the driver, but, in their current form, are not yet capable of providing self-driving experiences that are complete and cost-competitive. Their limitations include:

- The perception of the environment: so far, the fusion of available sensors and artificial intelligence is not capable of "seeing" and understanding the vehicle's surroundings as accurately as a human being can. Humans use a combination of learning and sensory inputs to detect and interpret events

as they occur, and anticipate likely scenarios. Thus, the great challenge is, as of now, to find compact technologies that make it possible to provide a 360-degrees perception of the environment with the minimum amount of computational resources.

- The cost: creating a 360-degree view of the vehicle's environment requires a combination of sensors and may cost more than consumers are willing to pay. Lidar-based systems provide 360-degree imaging but are complex, expensive, and not yet ready for the market (the lidar system used in the Google car, for example, costs about $70000).

Effective, alternative technological solutions could be found by looking at the nature, and in particular at human and animal vision, as it will be discussed in the following sections.

1.6 The vision in humans and animals

Perception is essential for both humans and animals in order to navigate in unknown environments and perform vital tasks. Among all the embodied sensors which humans and animals rely on, the eyes are often the most important ones as they are passive sensors providing very rich information on both the environment and the ego-motion. In fact, as already mentioned in section 1.3, passive sensors are more energetically efficient than active sensors, and are therefore preferred by nature. In addition, light provides probably the most rich information to be passively sensed in our world, making humans and most animals develop this sensory modality more than others. In particular, vision is the most important sense for flying animals, such as insects and birds, since good eyesight is essential for safe flight. Thereby, these animals present some features, such as visual acuity and motion detection, superior to that of other species.

As for every sensory system, the visual signals has first to be sensed and then to be processed in order to perceive useful information. Here, we will first present two of the most interesting and important features of both human and animal visual systems: light adaptation and motion perception. Then, we will focus only on the insects' vision, and in particular the fly's visual system, as it represents probably the most efficient visual system for performing motion detection with the minimum amount of computational resources.

1.6.1 Light adaptation

Animals, like humans, can easily navigate in indoor and outdoor environments where the light intensity can vary temporally within a 7-decade range, from night time with a cloudy sky ($\sim 0.01\,Lux$) to day time with full sunshine ($\sim 100000\,Lux$), and spatially up to 3 decades, from shaded to lightened regions in the same scene. To do this, their visual systems need to auto-adapt to the temporal and local average luminosity in order to keep good contrast sensitivity while "coding" luminance information within a limited electrical range.

Thus, light adaptation in human and animal retinas has been extensively studied since the early 1950s, using both intra and extracellular methods [36–43]. In all these studies, the relationship between light stimuli and photoreceptor responses has been documented, both in the dark and with background illumination, via an adaptation process described by the so-called Michaelis-Menten equation [44]:

$$V = V_m \frac{I^n}{I^n + \sigma^n}, \tag{1.1}$$

where V stands for the photoreceptor's response and V_m is its maximum value; I denotes the light intensity and n usually ranges from 0.7 to 1; σ is the adaptation parameter, corresponding to the light intensity giving half of the maximum response.

The first micro-electrode recordings of rod and cone responses were obtained on saltwater fish (*Gerridae*) by Svaetichin in 1953 [45]. In his pioneering study, Svaetichin discovered the S-potentials, as they were subsequently called by Oikawa *et al.* in [46], which stands for "slow potentials", referring to the slow adaptation process which occurs in the photoreceptor potentials when they are exposed to flash lights against a steady background.

However, the first mathematical description of the cone response given by equation (1.1) was provided by Naka and Rushton in the case of the freshwater fish (*Cyprinidae*) [36]. The equation (1.1) where $n = 1$ is therefore also known as the Naka-Rushton law. The same model was subsequently validated and applied to turtles' cones by Baylor *et al.* [37] and to monkeys' cones by Boynton and Whitten [38], who introduced the exponent $n < 1$ for the first time. Many studies were then carried out on vertebrates and invertebrates, all confirming the equation in (1.1) with various values of n and sometimes with different interpretations of the adaptation parameter σ (in the salamander [39], gecko [40], frog [41], locust, fly and dragonfly [42], for instance, and in the human fovea [43]).

Figure 1.8(a) shows the responses of dark- and light-adapted red cone photoreceptors recorded intracellularly in the retina of the turtle (*Pseduemys Scripta Elegans*) by Normann and Perlman [47]. As can be seen from this figure, the function $V(I)$ defined in (1.1) gives rise in the $Log(I)$ domain to curves with a fairly smooth "S" shape (continuous curves), where the slope of the "S" is given by the value of n ($n = 1$ in that case) and the lateral shift by the value of σ. For flying insects, a similar sensory adaptation mechanism capable of compensating for large changes in light intensity right at the photoreceptor level has also been discovered [48, 49] (Fig. 1.8(b)).

Based on the S-shaped curves shown in Fig. 1.8, two main features of the light-adaptation behavior can be described by the incident-light model [50]:

- as the background lighting changes, the entire S-shaped curve shifts along the light intensity axis, which corresponds to a change in the sensitivity of the photoreceptor in the neighborhood of the background light. In fact, after reaching a peak value caused by an increase/decrease in the intensity of the light (data points), the potential V gradually returns to a steady-state value, reflecting its adaptation to the background. This decrease/increase in V corresponds to a "slow" increase/decrease in the parameter σ (see equation (1.1));

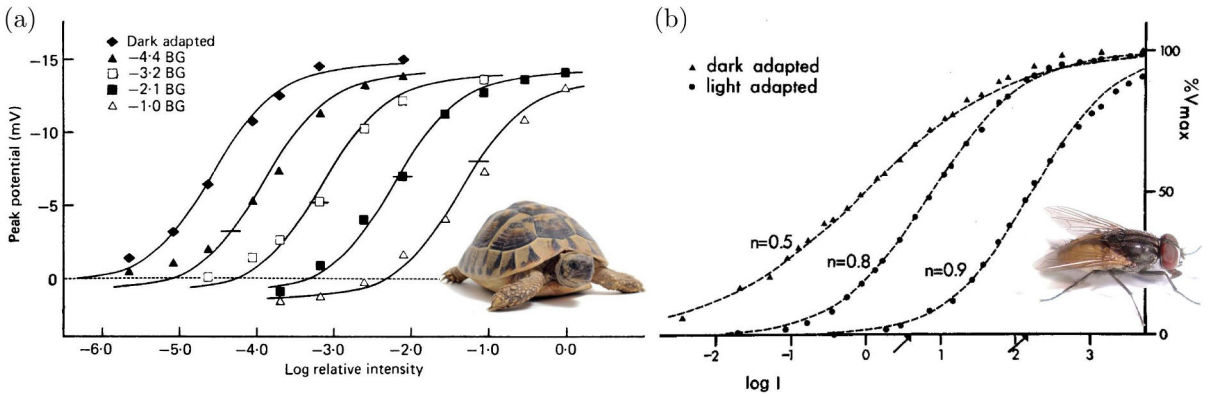

Figure 1.8: S-shaped curves corresponding to dark- and light-adapted response curves recorded in (a) a red cone of the *Pseduemys Scripta Elegans* turtle and (b) R1-6 photoreceptors of the *Calliphora* fly. Adapted from [47] and [48], respectively.

- as the background illumination increases, the operating point of the photoreceptor increases correspondingly (small horizontal lines in Fig. 3.1), which means that because of the non-linearity of the curve, the response to a given increment/decrement in the stimulus becomes smaller/larger at higher background levels. This process known as "response compression" was introduced for the first time by Boynton and Whitten in [38]. The slope of the curve around the operating point defines the contrast sensitivity.

1.6.2 Motion perception

Although it might appears as a straightforward process, motion perception is one of the most difficult problems to be explained in the vision process from both a psychophysical and physiological perspective.

In fact, two kinds of motion perception can occur when, for instance, two or more visual stimuli are switched on and off in alternation, depending on the stimulated area and the alternation rate: "first-order" and "second-order" motion perception, also known in psychophysics as "ϕ-phenomenon" and "β-movement", respectively [51].

At fast alternation rates and small stimulated areas, we obtain a first-order motion perception, also called "pure" motion perception since it is uncontaminated by form cues. This perception is mediated by relatively simple "motion sensors" in the visual system, which detect a change in illuminance at one photoreceptor in the retina and correlate it with a change at a neighboring photoreceptor after a short delay (see section 1.6.4).

On the contrary, by alternating the stimuli at relatively low rate and covering a relatively large region, we obtain a second-order motion perception, also called "apparent" motion since an object is perceived as moving when, in fact, a series of stationary images is being presented. This perception depends on the form of a moving contour, which is defined by contrast, texture or some other quality that does not necessarily result in an increase in illuminance or motion energy in the Fourier domain [52, 53]. There

is much evidence to suggest that early processing of first- and second-order motion is carried out by separate pathways before being combined in the V5/MT area of the human visual system [54].

The motion direction of a contour, e.g. a straight line, can be ambiguous for first-order motion detectors because the motion component parallel to the line can not be inferred based on early-stage visual inputs. This is because the neurons in the visual system which are responsible of such motion detection are sensitive to visual inputs in a small area of the visual field, as if each neuron was looking at the visual scene through a small aperture. This means that a variety of contours of different orientations moving at different speeds can cause identical responses in these neurons, hence the need of integrating this information with second-order motion perception.

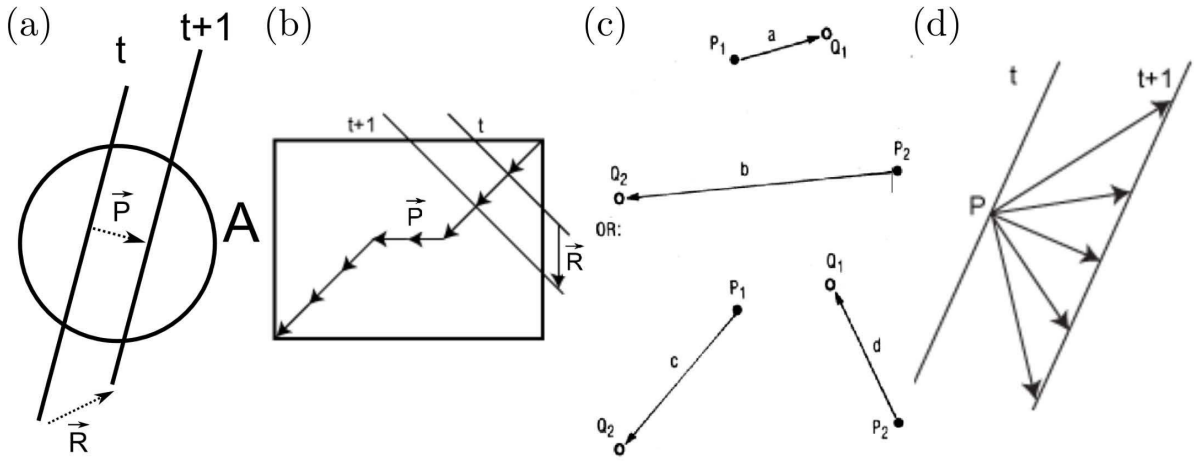

Figure 1.9: (a) The aperture problem. The real direction of motion of a one-dimensional object cannot be recovered when we look at this object through an aperture (receptive field) $\boldsymbol{A}$ smaller than the object. In this example, the real motion of the object is described by the vector $\boldsymbol{R}$ but the perceived motion $\boldsymbol{P}$ is only the component normal to the object. To determine the motion completely, a second stage must combine several local measurements. (b) The perceived motion through an aperture varies with the geometry of the aperture itself. In this example, the line appears to move diagonally, according to the normal direction of the line, only before reaching the corner of the aperture. After that, the line is perceived as moving horizontally (central part) because the borders of the aperture are parallel. (c) The correspondence problem. P_1 and P_2 are shown on the first frame, Q_1 and Q_2 on the next. Two one-to-one matches are possible ($P_1 \rightarrow Q_1$ and $P_2 \rightarrow Q_2$, or $P_1 \rightarrow Q_2$ and $P_2 \rightarrow Q_1$), leading to two possible patterns of motion. (d) A point on a line can have various motions since the point-correspondence appears ambiguous. Therefore, when the correspondence problem can not be solved because, for instance, of an aperture, humans perceive a motion normal to the edge. Adapted from [55] and [56].

This is also known as the "aperture problem" as described by Wallach [57]: in most cases, motion seen through a small aperture (receptive field) is perceived as normal to the edge of a line or a local surface contour (Fig. 1.9(a),(b)). From a second-order perception perspective, the aperture problem is closely related to the "correspondence problem", which refers to the capability of ascertaining which elements in

one image correspond to which elements in the previous image, where differences are due to the relative motion between the eye/camera and the objects in the scene (Fig. 1.9(c),(d)).

These two problems represent the main issues that computer-vision algorithms have to deal with in order to compute the right optical flow (see section 1.7).

1.6.3 The fly's visual system

The compound eye of insects is very different from all other animals' eyes, such as the mammalian eye (Fig. 1.10(a)). In fact, while in most animals' eyes the light is captured and projected on a layer of photoreceptors, i.e. the retina, through a single lens, in the compound eye the light is captured by a set repeating units called ommatidia, each consisting of a hexagonal lens, a transparent crystalline cone and a few photoreceptors, also called rhabdomeres [58] (Fig. 1.10(b)). In particular, the ommatidium of the fly comprises 8 photoreceptors, called R1-R8, which can be divided in 2 groups [59, 60]: two "inner" photoreceptors (R7 and R8) that participate in color vision with four different spectral sensitivities (from red to yellow), and six "outer" photoreceptors (R1-R6) that participate in motion detection and, in particular, are responsible for the optomotor response (see section 1.6.4).

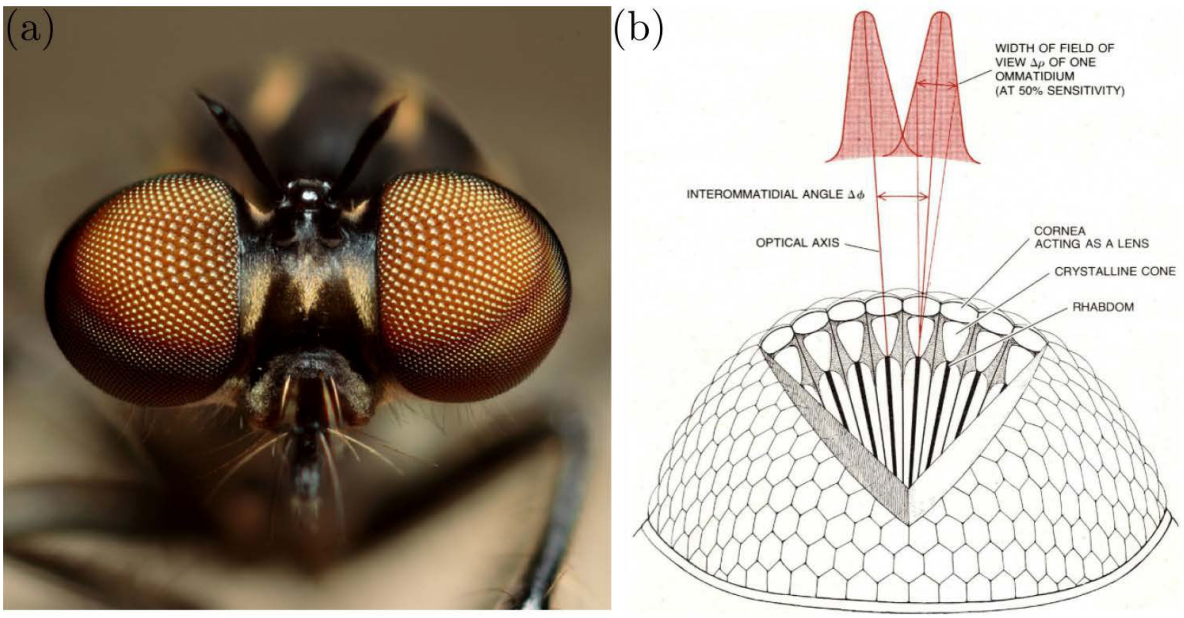

Figure 1.10: (a) Picture of the compound eyes of a Robber Fly (*Holcocephala fusca*) provided by G. Mazzarollo on `www.flickr.com`. (b) Simplified scheme of the compound eye of flying insects. Each elementary unit, called ommatidium, is composed of a hexagonal lens, a transparent crystalline cone and a few photoreceptors, also called rhabdomeres (here simplified to one). The ommatidia are characterized by their inter-ommatidial angle $\Delta\phi$ defining the spatial resolution, which varies along the compound eye. Due to the diffraction of light waves in the rhabdomere, each ommatidium features a Gaussian angular sensitivity characterized by its full width at half maximum $\Delta\rho$. Adapted from [58].

The ommatidia are distributed over the compound eye so that they form with their neighbors an angle $\Delta\phi$, called inter-ommatidial angle, which varies all along the compound eye defining its resolution. The smaller the inter-ommatidial angle, as, for instance, in the frontal region, the further objects can be detected, such as flowers, preys or predators [61]. Due to the small size of the lens (a few μm), the light is diffracted by the lens giving each ommatidium a Gaussian angular sensitivity characterized by its full width at half maximum $\Delta\rho$, also called acceptance angle [58]. Such an angular sensitivity, characterizing the field-of-view of each ommatidium, produces a spatial low-pass filtering of the visual scene which is essential in the motion detection processing [62]. Although the number of ommatidia as well as the inter-ommatidial angle ($\Delta\phi$) and the acceptance angle ($\Delta\rho$) largely vary among all flying insects, it was shown that in diurnal insects $\Delta\phi$ is roughly equal to $\Delta\rho$, ensuring that there is little or no aliasing without oversampling the visual information [61].

Two main types of compound eye can be found, depending on how the light converges onto the retina: the focal apposition compound eye and the refracting superposition compound eye [63] (Fig. 1.11(a),(b)). In the apposition eye, typically found in day-active insects (e.g. locusts and beetles), the ommatidia are optically isolated, each being sensitive to light along its optical axis. On the contrary, in the superposition eye, typically found in crepuscular or night-active insects (e.g. butterflies), the ommatidia are not optically isolated and therefore the light coming from one direction but passing through neighboring ommatidia converges into one photoreceptor, thus amplifying its response. A particular case of superposition eye is the neural superposition eye found, for instance, in diurnal flies, where the ommatidia are optically isolated (as in the apposition eye) but neuronal arrangement causes partial summation of the photoreceptors' responses (see [64, 65] for a review).

The main advantage of the compound eye is its compactness, allowing visual processing mechanisms to be situated immediately beneath the ommatidia, thus saving space and weight compared to the vertebrates' eyes where there is a wide useless space between the lens and the retina. In fact, beneath the ommatidia there are three optic lobes, called ganglia (or neuropils), namely the lamina, the medulla, and the lobula complex (lobula and lobula plate), corresponding to three different steps in the fly's visual processing [66] (Fig. 1.11(c)). The neurons in the first optic lobe, the lamina, receive directly inputs from the R1-R8 photoreceptors and respond as high-pass filters by amplifying temporal changes and therefore performing auto-adaptation to the background illuminance [48] (see section 1.6.1). These high-pass filtered signals, which may correspond to moving contrasts, are then transmitted to the medulla neurons to perform local motion detection between adjacent photoreceptors (see section 1.6.4). However, due to their small size, the medulla neurons are extremely difficult to record and little is known about them [67]. Lastly, the spatial convergence of local motion detectors is performed in the lobula plate where the information from several thousand of photoreceptors converge onto large neurons which are sensitive to specific motion directions.

The lobula plate is composed of approximately 50 tangential interneurons called Lobula Plate Tangential Cells (LPTCs) that receive inputs from local motion detectors over a wide region of the field of

view [68,69]. Among these neurons are the groups of horizontal (HS) cells and vertical (VS) cells, which are sensitive to specific directions of motion (e.g. from back to front for the H1 neuron) [70, 71]. Similar neurons have been discovered in the honeybee, called velocity-tuned (VT) cells and optomotor cells, which are sensitive to translatory and rotationary motions, respectively [72, 73].

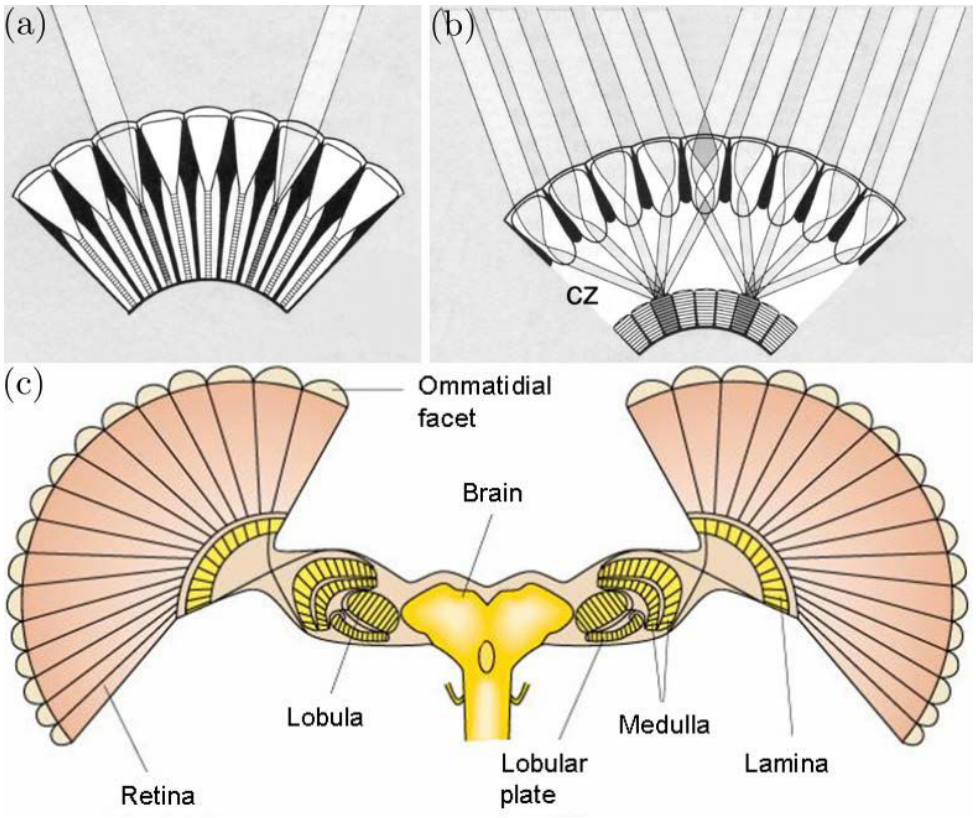

Figure 1.11: (a) A focal apposition compound eye. Light to photoreceptors comes through small corneal lens in each small eyelet. (b) A refracting superposition compound eye. A number of optical elements focus light to photoreceptors in the retina ("cz" stands for "clear zone"). (c) Schematic structure of the insect compound eye. The size and detailed structure of the different neuronal ganglia and centers may vary from species to species. The structure shown is closest to dipteran flies, although the number of retinotopic elements (facets and corresponding parts in deeper structures) is normally much larger. Adapted from [63] and [74].

1.6.4 Elementary motion detectors

Physiological and behavioral studies on animals and humans have shown that motion perception occurs in the early-stage processing of the visual system at higher frequency than what is meant to be the image rate. Various bio-inspired model, also known as elementary motion detectors (EMDs), based on these studies have been proposed for visual motion perception. Most of these motion detectors are based on a correlation (or de-correlation) mechanism which has been identified, for instance, in the medulla and

lobula lobes of the fly's visual system (see section 1.6.3), as well as in the low-level layers of neurons behind the human retina, namely the Outer and Inner Plexiform Layers (IPL and OPL).

The first and most famous correlation-based EMD was proposed by Hassenstein and Reichardt in the 1950s and is therefore known as the Hassenstein-Reichardt (HR) model. The HR model was initially inspired by electrophysiological experiments carried out in the visual system of the *Chlorophanus* beetle [75, 76], but was later shown to be consistent with the motion perception occurring in many other animals, such as other insects and particularly the fly [77], and even humans [78].

Correlation-based schemes compare the outputs of two photodetectors to illuminance variations at two adjacent positions. In particular, the original HR model was composed of 2 symmetric elementary units in which the output signal of the one of the 2 photoreceptors was multiplied by that of the other photoreceptor after being delayed by a time τ, which was generated by a low-pass filter (Fig. 1.12(a)). The output of the EMD is then given by the difference of the two units' outputs, which can be seen as the result of a correlation operation: a visual pattern moving to the right, for instance, will produce a high positive output whereas it will produce a high negative output when moving to the left. In particular, the highest positive (negative) value is obtained when the rightward (leftward) motion produces a time lag between the 2 photoreceptors' signals of exactly τ.

Improved versions of the HR model, also called elaborated Reichardt detectors (ERDs), have been proposed subsequently based on further biological findings, including different filtering blocks in the signals pathways, but without changing the essential correlation operation [80] (Fig. 1.12(b)).

In the 1980s, Franceschini et al. proposed a "facilitate and sample" model, which fits with an ERD but unveiled essential details about its inner processing structure, after electrophyisiologically recording the activity of the H1 neuron, one of the 50 wide-field LPTCs in the fly which is most sensitive to horizontal back-to-front motion (see section 1.6.3). In their experiments, instead of stimulating the eye by presenting some moving contrast in front of it, as it was done, for instance, by Hassenstein and Reichardt, they created an "apparent motion" in the field of view of one ommatidium by stimulating only 2 adjacent photoreceptors of the same ommatidium, in particular the R1 and R6 photoreceptors, thanks to a custom-made instrument hybrid between a microscope and a telescope [60, 62] (Fig. 1.13(a)). This new technique made it possible to isolate a single EMD and therefore reveal its inner processing structure in a more direct way than previous experiments.

The results obtained first showed that the H1 neuron firing rate increased when presented with a motion in the preferred direction (back-to-front) and was inhibited when the motion was in the non-preferred direction (front-to-back) [81] (Fig. 1.13(b)). This confirmed that the fly's EMD was composed of a slow "facilitating arm" which drives a parameter (e.g., a gain or a threshold) that adjusts the transfer of a fast "facilitated arm" (left and right side of the scheme in Fig. 1.13(d), respectively). The facilitating arm dynamics was identified as a second-order low-pass filter ($\tau \approx 50\,ms$) by measuring its impulse response when varying the interstimulus interval of short-lasting (10-ms) stimuli [62] (Fig. 1.13(e)). On the contrary, the facilitated arm dynamics was identified as a first-order high-pass filter ($\tau \approx 100\,ms$)

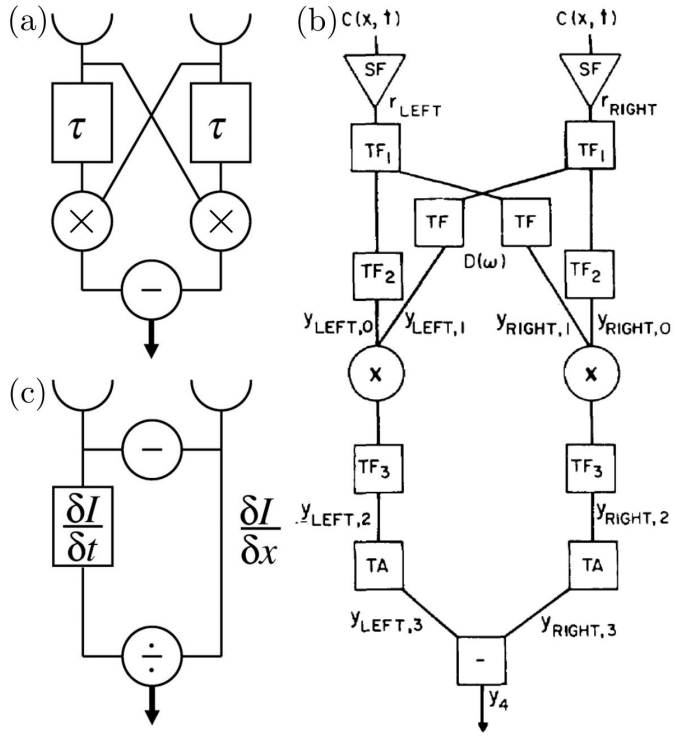

Figure 1.12: (a) Schematic representation of the original Hassenstein-Reichardt (HR) model. In this basic version of a correlation-based elementary motion detector (EMD), the input signal of one branch of each unit is delayed by a time interval τ and then multiplied by the non-delayed input signal of the neighboring branch. The final detector output is given by the difference of the 2 units' outputs. (b) Schematic representation of an elaborated Reichardt detector (ERD). The input is a luminance pattern with contrast $c(x,t)$; it is sampled by linear spatial filters (receptive fields, SF's) with spatial responses r_{left} and r_{right} centered at locations x_{left} and x_{right}; $y_{i,H}$ $(H = left, right)$ represents the signal at the various stages i for the left and right subunits. TF indicates a linear, time-invariant filter with Fourier transform $D(\omega)$, $\boldsymbol{X}$ indicates a multiplication unit, TA indicates a temporal integration operation, and $-$ indicates a unit that subtracts its left from its right input. (c) Schematic representation of a gradient-based detector. The temporal luminance gradient one photoreceptor's signal ($\frac{\partial I}{\partial t}$ on the left branch) is divided by the spatial luminance gradient between the 2 photoreceptors' signals ($\frac{\partial I}{\partial x}$ on the right branch). Note that the spatial gradient is approximated as the ratio between the difference of the 2 photoreceptors' signals ($-$ sign at the top) and the spatial distance of the photoreceptors' positions (Δx, missing in this representation). Adapted from [79] and [80].

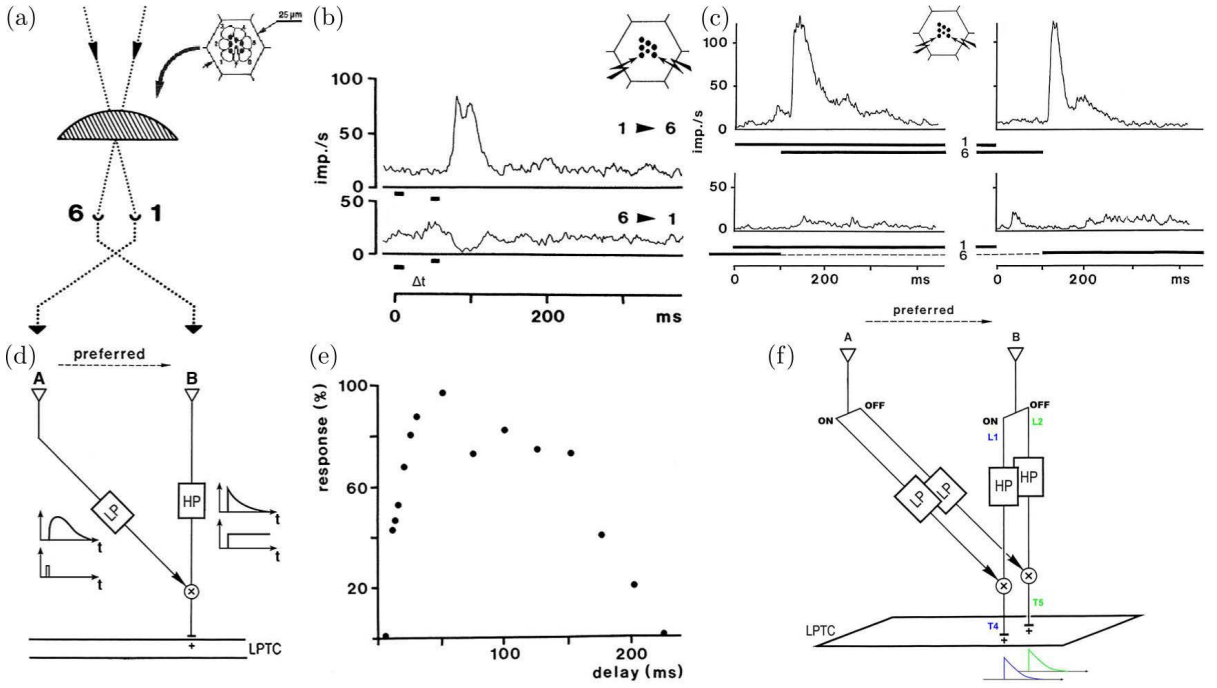

Figure 1.13: (a) Optical method used to drive the two input cartridges A and B of an EMD via single photoreceptor illumination. Sequential illumination of the photoreceptors R1→R6 drives the two cartridges A→B, producing "apparent motion" in the preferred direction, which conveys an excitatory signal (+) to the LPTC. (b) Response of H1 (spike instantaneous frequency) to a pair of brief (10-ms) flashes presented sequentially to the photoreceptors R1 and R6 (diameter $\approx 1\mu m$) of an ommatidium (inset) (spot diameter $\approx 1\mu m$, interstimulus interval $\Delta t = 50\,ms$, repetition time $= 500\,ms$, average $n = 100$). Note that the sequence simulating motion in the antipreferred direction (R6→R1) inhibits H1: this suggests the existence of a coextensive, mirror-symmetric EMD half-detector driven by the same two cartridges B and A and sensing motion in the opposite direction. (c) H1 responses to mixed sequences of long-lasting ($2\,s$) light and dark pulses received by two single photoreceptors R1 and R6. Among the eight step sequences that can be produced by manipulating the sign and the delay of the changes of brightness delivered to R1 and R6, only the two sequences shown in the two upper plots triggered a clear-cut response in H1. (d) Block diagram of the elaborated HR half-detector (sensing motion in the preferred direction) resulting from micro-stimulation experiments. The box in the left arm describes a second-order low-pass (LP) filter, the impulse response of which was measured in (e). The box in the right arm describes a first-order high-pass (HP) filter, the step response of which was measured in (c). (e) Impulse response of the low-pass filter of the lateral facilitating arm in (d), elicited by presenting R1 and R6 with sequences of brief (10-ms) flashes with various interstimulus intervals (data are means of H1 recordings on nine houseflies). Adapted from [60]

by measuring its step response when exposing the 2 photoreceptors to a sequence of long-lasting (2-s) stimuli [62] (Fig. 1.13(c)), which represented the first major difference with the original HR detector. As a result, the H1 neuron was not activated when the delay between the stimuli was too short ($< 10\,ms$) or too long ($> 230\,ms$) which confirmed the fact that the EMD responds only to motions within a fixed range (with different gains depending on the velocity), corresponding to angular velocities (i.e. the optic flow) between $16^o/s$ and $360^o/s$ for the considered ommatidium having an interreceptor angle $\Delta\varphi \approx 3.6^o$.

In addition, the appearance of two peaks in the H1 neuron's response when applying long-lasting ($1\,s$) light sequences without overlap, could be explained by the existence of two separate pathways sensitive respectively to ON-ON sequences (light-to-light transitions) and OFF-OFF sequences (dark-to-dark transitions) (Fig. 1.13(c)). This result suggested to double the pathways of the initially-proposed EMD giving 2 separate ON and OFF pathways (Fig. 1.13(f)), which represented another important difference with the original HR detector. Recent experiments have confirmed the existence of separate brightness increments and decrements circuits controlled by the L1 and L2 neurons in the lamina [82].

It is worth noting that ERDs have an important drawback: the output depends on not only on the velocity of a moving pattern but also on its spatial structure, which is obviously undesired for motion estimation. For this reason, other models of EMDs have been proposed throughout the years, although showing no consistency with the biological findings.

A prominent alternative model for motion detection is the so-called "gradient detector", which computes a velocity-dependent signal by dividing the temporal derivative of local luminance $\frac{\partial I}{\partial t}$ by its spatial derivative $\frac{\partial I}{\partial x}$ [83–86] (Fig. 1.12(c)). Differently from the ERD, the gradient detector provides a signal that is proportional to the image velocity at each point and does not depend on pattern properties. In particular, no modulations are expected in the local signals as long as the velocity is constant, and the velocity dependence of the global signal should not vary with the spatial wavelength of the pattern. The main issue with this computational model is that the uncertainty in the computed velocity may vary widely. In fact, if the spatial derivative $\frac{\partial I}{\partial x}$ is small, the noise in the temporal derivative is amplified, and eventually when the spatial derivative is close to 0, the velocity is completely undefined. Along these lines, Potters and Bialek proposed that an ideal motion detector scheme would be based on the gradient detector only for high signal-to-noise, whereas it would be an ERD at low signal-to-noise ratios [79, 87].

However, experiments on the fly visual systems provided unambiguous evidence in favour of the Reichardt detector under all luminance conditions [88]. This "suboptimal" computation could be explained by the fact that Reichardt detectors have an automatic gain control allowing them to dynamically adjust their input-output relationships to the statistical range of velocities presented, while gradient detectors do not have this property. As a consequence, Reichardt detectors, but not gradient detectors, always provide a maximum amount of information about stimulus velocity over a large range of velocities [79].

Interestingly, the most famous methods for computing the optic flow from image sequences are closely related to gradient detectors instead of Reichardt detectors (see section 1.7.2), going (apparently) in the opposite direction of nature.

1.7 The optical flow

Optical (or optic) flow can be defined as the projection on the retina of the (apparent) motion of visual contrasts (objects, surfaces, edges, etc) caused by the relative motion between an observer (eye or camera) and the environment [89] (Fig. 1.14). In other words, the optic flow is given by the change of structured light in the image, i.e. on the retina of an eye or a camera, due to this relative motion between the observer and the scene. In mathematical terms, the optic flow is an angular velocity vector field, which is generally complex as it results from the motion of an agent relatively to moving or stationary objects in a 3-D unstructured, unknown environment projected in a 2-D image. Further definitions can be found in the literature highlighting different properties of optic flow.

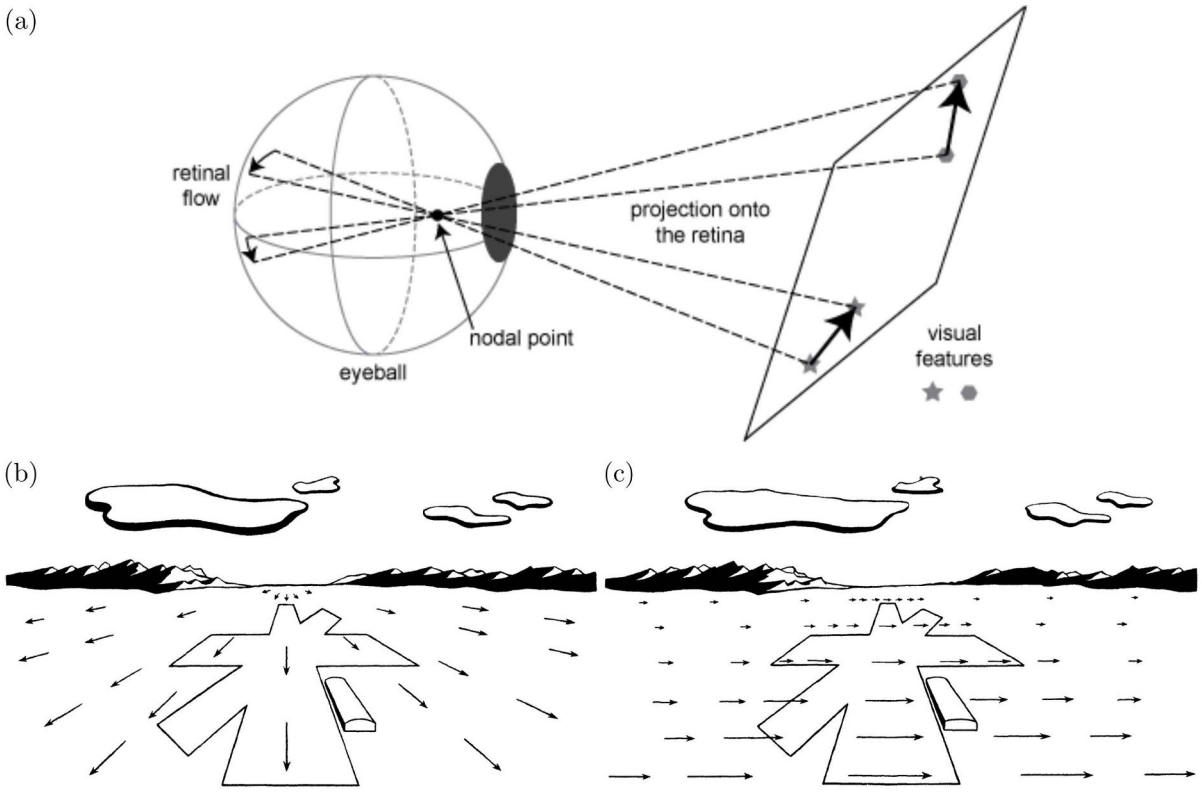

Figure 1.14: (a) The optic flow is generated on the retina by changes in the patterns of light. This example shows the movement of two visual features (star and hexagon) on a plane and their angular displacements on the retina. (b),(c) Schematic representation of the optic flow vector field generated by the apparent visual motion of the ground on the retina of a flying agent when it is moving (b) forwards and (c) leftwards. Each vector represents the displacement vector/projection of the angular velocity of a point of the environment from one image to the next one. In (b), the point right in front of the agent where the angular speed is zero is called the focus of expansion (FOE). Adapted from [56] and [89].

The first studies on how humans and animals use this visual motion information for perceiving the environment was carried out in the early 1900s by the physician Hermann von Helmholtz. He was one

of the first scientist who distinguished from both a physiological and psychological point of view what is "visual sensation" (color, contrast, adaptation, etc) and "visual perception" (depth, motion, etc) [90]. He stressed for the first time the importance of motion perception, instead of colors and edges perception, for recognizing the objects' shapes, and in particular for one-eyed persons: "My belief too is that it is mainly by variations of the retinal image due to bodily movements that one-eyed persons are able to form correct apperceptions of the material shapes of their surroundings" (from [90], p. 297).

However, the concept of optical flow was introduced by the psychologist James J. Gibson in 1950 to describe the visual information that makes animals navigate into the environment [89]. Since Gibson, authors have further demonstrated the role of the optical flow for the perception of the shape, distance and movement of objects, as well as the control of locomotion [91–94], and have also formulated various analytic models [95–97].

Although the phenomenology of the optic flow has been qualitatively and quantitatively known for decades since von Helmholtz, only during the early 1980s popular algorithms were proposed for the estimation of optic flow from image sequences [98, 99]. In this context, the optic flow field depends not only the relative motion and the geometry of the scene, but also by the geometry of the pinhole camera, namely the projection function.

1.7.1 Modeling of OF

In their analytic model, Koenderink and van Doorn described the optic flow (or motion parallax) $\hat{\boldsymbol{v}}$ of a fixed "fiducial point" P (e.g. a landmark) with respect to a moving "vantage point" O, corresponding to the nodal point of the retina, as follows [100]:

$$\hat{\boldsymbol{v}} = \frac{1}{||\boldsymbol{p}||}((\boldsymbol{T} \cdot \hat{\boldsymbol{p}})\hat{\boldsymbol{p}} - \boldsymbol{T}) - \boldsymbol{R} \times \hat{\boldsymbol{p}}, \tag{1.2}$$

where $\boldsymbol{p}$ is the position vector of P with respect to O, $\hat{\boldsymbol{p}} = \frac{\boldsymbol{p}}{||\boldsymbol{p}||}$ is the unit vector representing the direction of $\boldsymbol{p}$, and $\boldsymbol{T}, \boldsymbol{R}$ are the translational and rotational vectors corresponding to the motion of the vantage point with respect to an inertial frame (Fig. 1.15).

We note that, although the motion parallax is originally a 3-D vector field ($\boldsymbol{v}$), the resulting optic flow field ($\hat{\boldsymbol{v}}$) can be described as a 2-D vector field as it is always orthogonal to $\hat{\boldsymbol{p}}$. Thus, the optic flow is usually represented using the Mercator projection defined by the azimuth (Φ) and elevation (Θ) angles.

By looking at equation (1.2), we can first note that the optical flow produced by self-motion of an observer that is both translating and rotating can be written as a linear combination of the optical flow vectors that would have been produced by separate translation and rotation (Fig. 1.15). Then, we note that only the translational component of the optic flow field depends on the distance of the point P, and this is why animals and humans usually try to reduce any rotational component of their head in order to stabilize their locomotion using optic flow cues (see section 1.7.3). In fact, if the observer's rotation ($\boldsymbol{R}$) is null (or known), the distance to P ($\boldsymbol{p}$) or the observer's translation ($\boldsymbol{T}$) can be computed when one of these two is known. However, if P is also moving with a velocity $\boldsymbol{T}_p$ with respect to the inertial frame,

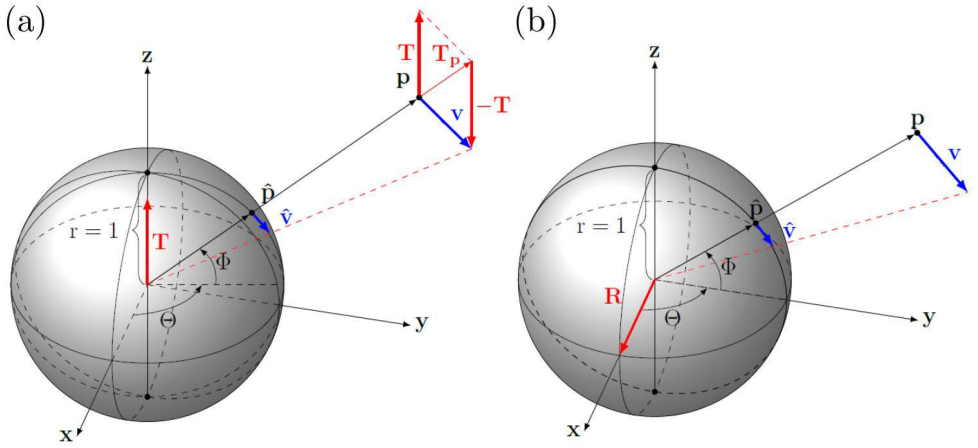

Figure 1.15: (a) The translation T of a spherical sensor (gray sphere) causes a relative motion of the point p in the sensor's environment that leads to the movement of the point's projection $\hat{p}$ on the sensor's surface. The optical flow vector $\hat{v}$ describes the velocity of the projected point $\hat{p}$. (b) The rotation R of a spherical sensor (gray sphere) causes a relative motion of the point p in the sensor's environment with an angular velocity v. The optical flow vector $\hat{v}$ describes the velocity of the projection of the point $\hat{p}$ on the sensor's surface. Adapted from [101].

then $\boldsymbol{T}$ must be substituted with $\Delta\boldsymbol{T} = \boldsymbol{T} - \boldsymbol{T}_p$ in equation (1.2). In this case, e.g. in dynamic, unknown environments, additional information on either $\boldsymbol{p}$ or $\boldsymbol{T}_p$ must be provided to solve equation (1.2).

1.7.2 Computing the OF

A large number of methods for computing the optic flow from image sequences has been proposed throughout the years. Most of these methods can be classified in three main groups: gradient-based, block-matching and frequency-based methods.

Gradient-based (or differential) techniques compute the optic flow from spatio-temporal derivatives of the image intensities. The image domain is therefore assumed to be continuous and differentiable in space and time.

Block-matching techniques, compute the optic flow as displacement vectors of some identifiable blocks/features between subsequent images, using, for instance, cross-correlation techniques.

Frequency-based techniques compute the optic flow from the Fourier transform of the image sequence as the best-correlated displacements by applying, for instance, orientation-sensitive filters, such as the Gabor filter.

Lastly, bio-inspired (or event-based) methods based on quasi-continuous neighboring-pixels' signals, instead of fixed-rate image sequences, can be considered as a fourth separate group.

Gradient-based methods

Let us consider the scalar function $I(x, y, t)$ describing the intensity values of an image sequence from a camera, where x and y denote the spatial coordinates of each pixel whereas t denotes the time. Then, the optic flow can be defined as the dense vector field $\boldsymbol{v} = [u(x, y, t) \quad v(x, y, t)]^T \approx [\frac{\Delta x}{\Delta t} \quad \frac{\Delta y}{\Delta t}]^T$ describing the displacement $(\Delta x, \Delta y)$ of these intensity values between 2 subsequent image frames sampled by a time Δt. For simplifications, we assume that the image sequences are gray-value or reduced to one channel. More recent methods have been developed that take RGB images as input (e.g. in [102]).

Most gradient-based methods can be derived by considering that brightness remains nearly constant between 2 subsequent image frames, i.e. $I(x, y, t) \approx I(x+\Delta x, y+\Delta y, t+\Delta t)$. By considering a first-order approximation of the Taylor series of the latter, the following continuity equation can be derived:

$$\nabla I \cdot \boldsymbol{v} + \frac{\partial I}{\partial t} = 0, \tag{1.3}$$

where $\nabla = [\frac{\partial}{\partial x} \quad \frac{\partial}{\partial x}]^T$ is the spatial gradient operator.

Equation (1.3) has 2 unknowns (u, v) and therefore can not be solved as such but another set of equations is needed. This limitation can be seen as a computational consequence of the aperture problem (see section 1.6.2). Such equations can found by applying some additional local or global constraints to the optic flow field, determining therefore 2 subgroups, namely local and global differential methods.

The most popular local method is the Lucas-Kanade method [98], also known as Lucas-Kanade-Tomasi feature tracker for image-sequences registration, which assumes that the velocity vector $\boldsymbol{v}$ is nearly constant within a small neighborhood surrounding a pixel. After dividing the image in regions, equation (1.3) is calculated for the intensity values of the pixels in each region and the optic flow field is then computed by performing a least-squares minimization to find the best fit for $\boldsymbol{v}$.

For more robustness, a weighted version of the least-squares solution is usually performed to give more weight to the pixels closer to the central pixel of the each region, for instance, by considering a Gaussian function. Various methods performing outliers rejection (e.g. in [103]) or multiple motions estimation in a generalized structure tensor [104] have also been proposed. To increase the algorithm performance while decreasing the computational cost of its implementation, a pyramidal approach was presented in [105] and has been widely used in various robotic applications.

The most popular global method is the Horn-Schunk method [99], which uses variational calculus to minimize the functional obtained by adding a smoothness constraint on $\boldsymbol{v}$, i.e. $||\nabla u||^2 + ||\nabla v||^2$, to the right-hand term of equation (1.3). Various methods using different regularity constraints have also been proposed throughout the years [106, 107].

Local methods may offer relatively high robustness under noisy conditions, but may be unreliable if the optic flow is not uniform within each region and give rather sparse flow fields. On the other hand, global methods allows optic flow to vary smoothly within neighborhoods and yield dense flow fields, but have higher computational cost and are experimentally known to be more sensitive to noise. Thereby, more recent methods have been proposed to unify both methods into one [108].

As differential methods rely on the hypothesis that the image sequences are differentiable, these methods may suffer from undersampling or discontinuities in the optic flow when visual motion is too fast with respect to the frame rate or is produced by occluded objects.

Discontinuity in the optical flow due to occlusion can be analyzed by using line processes mixed with velocity distributions or parametric models [109]. These techniques perform the segmentation of optical flow into regions corresponding to various independently moving objects or surfaces.

The problem of undersampling may lead to inaccurate estimates of the optic flow at low and high velocity as the accuracy depends on the frame rate of the acquisition system. By adapting the frame intervals based on the speed of the objects, Chen et al. (2012) proposed a scheme based on the Lucas-Kanade algorithm and using high-frame-rate cameras, which was able to process optic flow in a wide range of velocities [110].

Although not exactly a gradient-based method, Srinivasan (1994) proposed a new algorithm called the Image Interpolation Algorithm (I2A), which minimizes the error at each time step between the acquired image and the interpolated one based on a set of the previous images [111]. The main advantage of this algorithm is its robustness in comparison to the required computational power which makes is suitable to be embedded on microcontrollers.

Block-matching methods

Numerical differentiation may sometimes lead to inaccurate or wrong estimates because of undersampling (frame rate too low compared to motion) or poor signal-to-noise ratio. In these cases, gradient-based methods are often inappropriate, whereas block-matching methods may give better results.

These approaches approximate the optic flow vectors by the displacements that yield the best fit between some blocks in the image sequence, which could be simply regions or identifiable features/patterns (objects, edges, surfaces, etc). The best fit is usually performed by maximizing a similarity measure, which is often represented by some correlation coefficients such as the Pearson cross-correlation coefficients. Other similarity measures, such as the absolute value of differences or the sum of squared differences, can be used to reduce the computational cost, which can be very high when performing cross-correlation, but the performances often decrease in these cases.

The main limitation of these methods is related to the correspondence problem (see section 1.6.2). In fact, the risk is to match two different blocks over time because they “look” similar, leading to a wrong estimation of the optic flow, and this risk increases when the frame rate is too low compared to the image velocity. Thereby, methods involving hierarchical or probabilistic approaches have also been proposed [109] to help make a robust estimation of the optic flow.

As opposed to differential methods, block-matching methods may deliver very sparse flow fields when only few features can be detected in the images, and usually require higher computational resources for performing image processing to extract features.

Frequency-based methods

Frequency-based methods apply orientation-sensitive or phase-based filters (e.g. Gabor filters) to the Fourier transform of image sequences to find the best-correlated displacements from an oriented or phase energy point of view [112, 113].

In some cases, these methods can correctly estimate the optic flow in image sequences for which both gradient-based and block-matching approaches would fail. For example, the motion of random dot patterns can be easily extracted from the resulting oriented energy in the Fourier domain [112], whereas it may be difficult to estimate with feature-based methods.

However, these methods have even higher computational cost then correlation-based methods and may yield ambiguous results, for instance, for periodic images (i.e. with similar contiguous regions) due to presence of several correlation peaks in the Fourier domain.

Although some of these models showed consistency with some psychophysic experiments, there is no evidence that they are actually performed in the animals or humans visual system.

Bio-inspired methods

As already mentioned at the beginning of this section, vision in animals is a very complex process which can not be simplified with a fixed-rate image processing as it happens in standard cameras. Therefore, methods inspired by the elementary motion detectors (EMDs) found in the animal and human visual systems (see section 1.6.4) have been presented during the last few decades for local optic flow estimation using signals from neighboring photoreceptors rather than image sequences [114, 115].

Although both hardware and software implementations of elaborated Reichardt detectors (ERDs) have been successfully presented for detecting motion and reproducing visio-motor responses on bio-inspired robots [115], these models have some obvious limitations. First, the output of ERDs does not represent exactly the optic flow but is somehow proportional to it in an undefined way, therefore the measurements obtained can not be used directly, for instance, for estimating velocities or distances. Second, the velocity of motion must lie within a certain range, determined by the delay (or the low-pass filter) and the interreceptor distance. A wide range of velocities can be covered by using several detectors with different internal delays and interreceptor distances, but such solutions would be computationally costly and therefore less appealing for robotic applications.

In addition, the direction of motion in 2-D cannot be determined reliably from the output of few neighboring detectors as they can not solve the aperture problem (see section 1.6.2). In order to solve it, the output of several correlators have to be integrated spatially over a larger field of view as presented, for instance, in [116], but this is once again computationally costly and therefore unsuitable for real time applications.

By taking inspiration from their electrophysiological experiments carried out on the fly's H1 neuron (see section 1.6.4), Franceschini et al. proposed a token-matching scheme [62, 117–119] which has been later called the "time of travel" (TOT) scheme [120, 121]. Differently to ERDs, TOT schemes made it

possible to directly estimate the time delay Δt between 2 neighboring photoreceptors' signals induced by a visual motion (Fig. 1.16). The 1-D optic flow ω can be then computed by diving the interreceptor angle $\Delta\varphi$ by this delay, as follows:

$$\omega = \frac{\Delta\varphi}{\Delta t}. \tag{1.4}$$

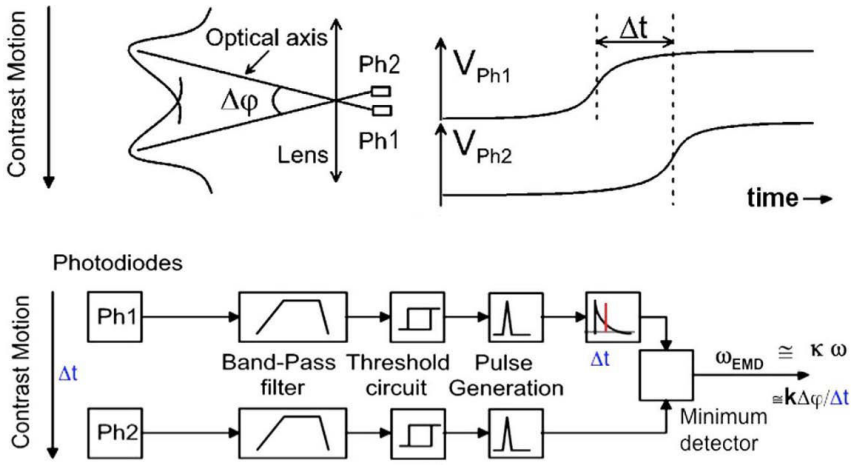

Figure 1.16: Principle of the time-of-travel EMD. The motion of a contrasting edge occurring in the FOV of a pair of photodiodes Ph1 and Ph2 (interreceptor angle $\Delta\varphi$) induces a time lag Δt in their electrical responses. After low-pass (spatial) filtering and bandpass (temporal) filtering steps, the signals from Ph1 and Ph2 were thresholded, generating a unit pulse in both pathways. The pulse from pathway (2) was used to sample the decaying exponential function generated by the pulse from pathway (1) by means of a minimum detector, the output of which therefore increased monotonically with the local OF $\omega = \frac{\Delta\varphi}{\Delta t}$. Adapted from [122] and [117].

Several versions of time-of-travel EMDs, also called Local Motion Sensors (LMSs), have been presented by our laboratory during the last 20 years, as it will described in details in section 1.8.3.

In [123], the authors suggested a correlation-based TOT scheme in which signals from neighboring pixels were delayed by different time delays and the cross-correlation between the delayed and non-delayed signals was computed in parallel. Then, the time delay Δt could be estimated as the time delay giving the maximum cross-correlation. However, no implementation of this method has been presented so far.

1.7.3 Using the OF

As described in section 1.7.1, optic flow can provide very rich information on both the environment and ego-motion while being processed at high frequency in the early stage of the visual system (see section 1.6.4). Therefore, optic flow is largely used in nature as well as in most of the technological applications involving visual sensors, whether they are bio-inspired or not.

Various biological studies have shown that humans and animals use somehow the optic flow generated

by their own motion to navigate in unknown environments [89,92,94,124–126]. This is particularly true for flying animals, such as insects and birds, since they can not rely on proprioceptive sensors to gauge, for instance, strides count and length like in legged animals.

Optic flow can contribute to navigation in different ways, depending on the animals, the tasks and the environmental conditions. Although little is known on how exactly humans and animals use optic flow, it is found to be used in all main functions of locomotion and flight control, such as gaze control, direction and speed control, altitude control and obstacle avoidance, and path integration (odometry).

Most of the knowledge we have about the use of optic flow is on insects, especially on bees, since they are relatively easy to manipulate and their "simple" neural processing make it possible to extract functional models from behavioral experiments with more certainty.

Compensating rotational movements

The stabilization to a straight flight is a very important feature for flying insects, as the rotational optic flow only contains information of the insect's own rotations, whereas the translational optic flow depends on the ratio between its horizontal speed and the distance to the objects (see section 1.7.1). This is the reason why mechanisms such as the optomotor response, the vestibulo-ocular reflex, the head reorientation and the saccades, have been found to help the insects to always perform translations in order to perceive less complex optic flow fields, hence facilitating extraction of useful information.

Compensation of yaw rotations have been found for the first time the 1950s by Hassenstein using a "Spangenglobus" ("Y-maze globe") [127], showing that insects react to visual stimuli produced by rotations by moving in the same direction of such rotations. This behavior, also known as the "optomotor response", was confirmed by subsequent experiments carried out by Franceschini in 1973 in which a head-fixed housefly was tethered to a custom micro "bicycle", while it was presented to specific optical stimuli that were not perturbed by the animal's body or head movements. The virtual trajectory produced by the fly's strides on the bicycle in response to a drifting grating presented in its frontal FOV clearly showed the reaction of the insect to maintain a straight course by compensating for undesired deviations [60] (Fig. 1.17). About 25 years later, Srinivasan et al. could characterize quantitatively this optomotor response by tethering an insect in a rotating striped drum, showing that the insect tried to turn in the direction of the rotation by producing a corresponding yaw torque [115]. Such a behavior has been successfully explained by implementing ERDs on wheeled robots (see section 1.9).

Similar reactions for compensating roll and pitch rotations, namely the vestibulo-ocular reflex and the head reorientation, have been found in insects, as well as in other animals and even humans, in order to keep a straight gaze and therefore extract more straightforward optic flow cues [128–130].

Controlling speed and direction

Inspired by observations on mosquitoes and locusts, Kennedy hypothesized as early as the 1940s that flying insects control their speed by maintaining a preferred retinal velocity with respect to the ground

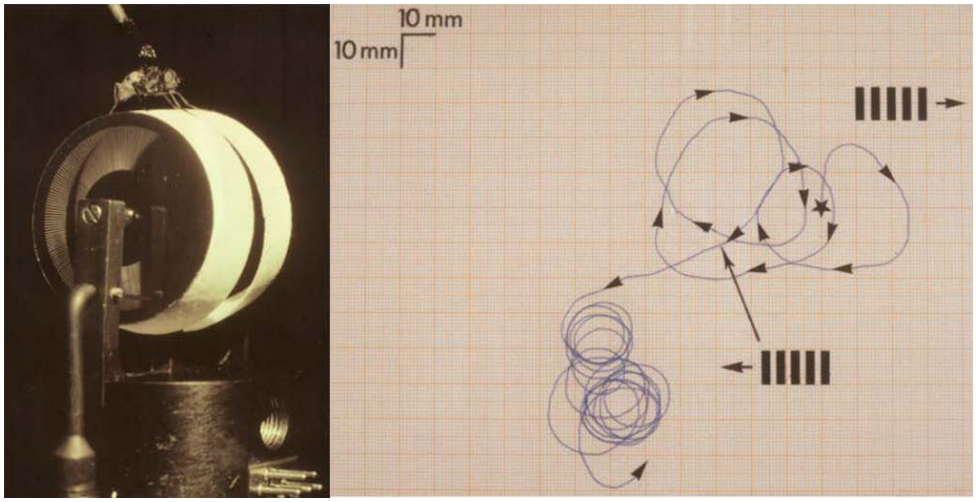

Figure 1.17: The Mark I Fly Bicycle (1973) constructed for analyzing insect's optomotor responses qualitatively. A head-fixed housefly was tethered to the micro bicycle while it was presented to visual gratings moving leftwards and rightwards. The virtual trajectory produced by the fly's strides on the bicycle in response to these moving gratings clearly showed the reaction of the insect to maintain a straight course by compensating for undesired deviations. Adapted from [60].

below them [131, 132]. Nearly 50 years later, experiments performed on honeybees have shown that the optic flow is used to control both their forward and lateral speed [92, 126, 133, 134]. In particular, it has been observed that honeybees flying into a narrow tunnel tend to (i) navigate at equal distance from the lateral walls [92] or follow one of the two walls at a certain distance [135] (depending on their initial position and on the texture of the walls), and (ii) adjust their speed depending on the size of the tunnel [92] in both the horizontal and the vertical planes [136] (Fig. 1.18). In addition, it has been shown that, in a tunnel with moving lateral walls, flight speed increases when the wall is moving in the direction of flight (decreasing the perceived optic flow) and vice-versa [126, 136]. These studies led to the conclusion that the bees adjusted their flight (speed and direction) by keeping constant optic flow cues. Recent experiments have shown that similar flight control strategies based on optic flow cues are also performed by birds [137].

In 2001, Warren et al. experimentally showed for the first time that humans use optic flow to control their walking [94], as it was already predicted by Gibson 50 years earlier (see section 1.7). In their experiments, the subjects walked freely in a room while wearing a stereoscopic head-mounted display that created an immersive virtual environment where the focus of expansion (FOE) was displaced of 10^o with respect to the theoretical one. The results showed that people steered to the target by following trajectories much closer to that predicted by the optic flow than that predicted by the egocentric visual perception (i.e. their position estimated with respect to the target) whenever the visual scene produces good-enough optic flow cues [94] (Fig. 1.19). These and subsequent results [138–140] have been particularly important to design and develop both static and dynamic virtual reality simulators used, for

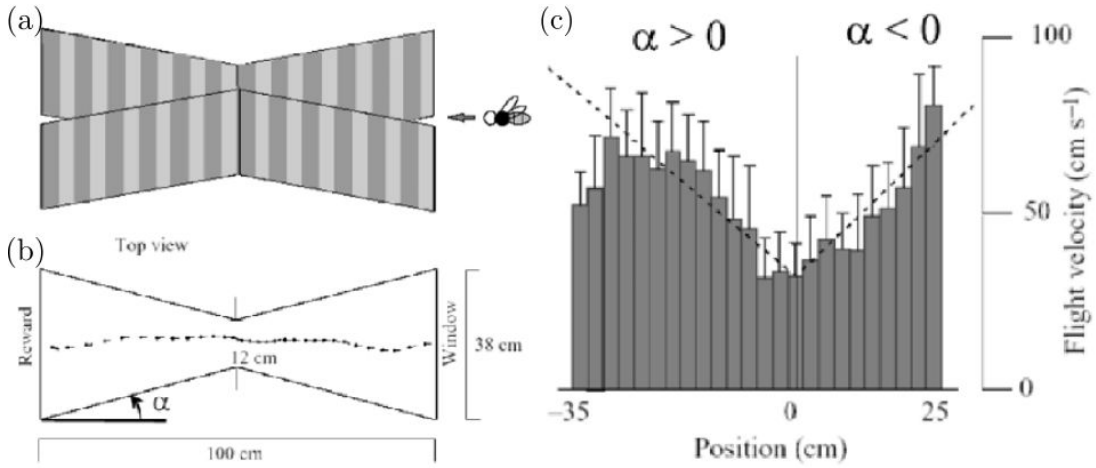

Figure 1.18: (a) Honeybees were trained to fly in a 1-m long tapered tunnel lined with vertical black and white gratings. The width of the tunnel varies from 38 *cm* to 12 *cm*. (b),(c) Flight speed decreases as the tunnel narrows, and increases as it widens. Adapted from [92].

instance, for automotive tests, that reproduce consistent visual perception of the environment [141,142].

From all these results, authors have proposed various optic flow-based strategies for controlling both flight and locomotion [115, 143], some of which have been successfully implemented on flying and wheeled robots (see section 1.9).

Estimating the distance traveled (Odometry)

Behavioral experiments have shown that insects most likely use optic flow information also to estimate their complex trajectories when, for instance, they are searching for new food sources, allowing them to take a direct route back to their homes. For example, after returning to their hives, the bees communicate the distance and direction of the food source to their nest mates by encoding the information in a "waggle dance" [144]. The length of this waggle run signals the distance flown and the orientation of the waggle axis (relatively to gravity) signals the azimuthal direction of the food source, relative to the direction of the sun [144]. Authors showed that these bees may estimate the distance flown by integrating the optic flow perceived during flight [145–147]. In fact, when moving the ground or the lateral walls forwards or backwards, the honeybees estimated respectively longer or shorter distances than that they were trained for, and this estimation did not directly depended on the textures used. However, it has been suggested that such an odometer would be influenced by the richness of visual information of the overflown terrain [148] as well as the presence of celestial cues [149].

Similar studies on desert ants showed that optic flow appears to contribute also to the desert ant odometer, although it mainly relies on strides count and length as for most of legged animals [93] (Fig. 1.20). This is true for the ventral part of the ant's eye, which is the part looking at the desert floor and thus suitable to monitor optic flow during locomotion, whereas differently to flying insects, lateral optic flow is apparently without influence on distance estimation [150, 151].

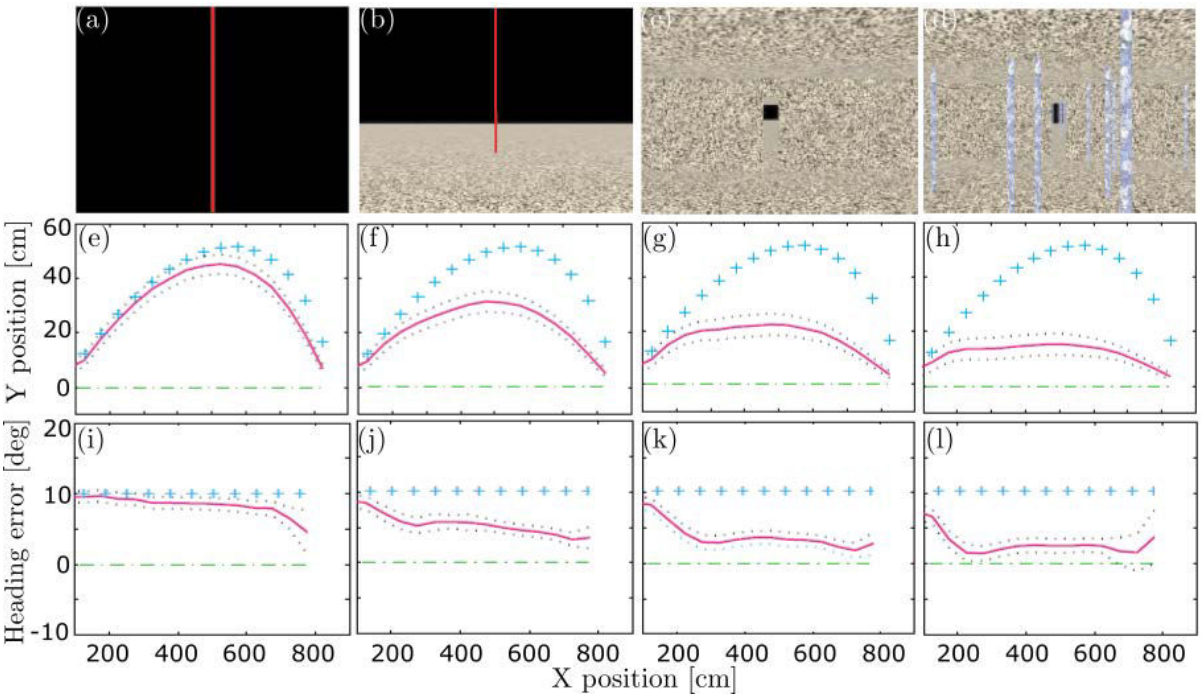

Figure 1.19: Experiments where the subjects walked freely in a room while wearing a stereoscopic head-mounted display that created an immersive virtual environment where the focus of expansion (FOE) was displaced of 10^o with respect to the theoretical one. (a)-(d) The four virtual worlds: target line, line with ground, doorway, doorway with posts. (e)-(h) Mean path in the virtual world when using the corresponding virtual environment above. (i)-(l) Mean virtual heading error with respect to the longitudinal position when using the corresponding virtual environment above. Data are collapsed and plotted as though the displacement (10^o) is to the right (dashed lines), together with the between-subject standard error (dotted lines). The predictions of the egocentric direction hypothesis and the optic flow hypothesis are indicated by the crosses and dashed-dotted lines, respectively. Adapted from [94].

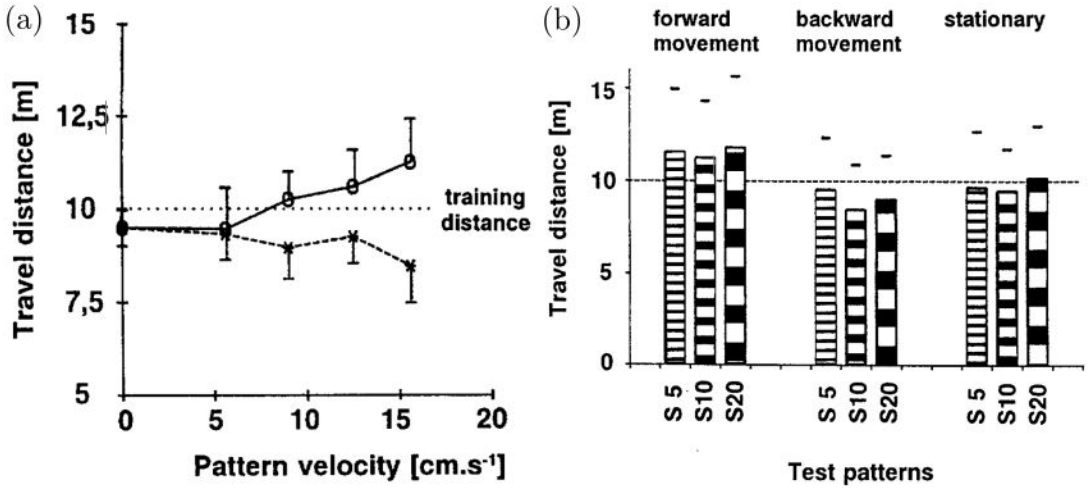

Figure 1.20: Experiments where desert ants were trained and tested to walk into a tunnel while moving the floor forwards and backwards. (a) Pattern movement in the forward direction (circles) leads to increased travel distances, while movement in the opposite direction (stars) decreases the lengths of the ant's paths. Training pattern: Sl0, stationary; test pattern: S10, moving. (b) Tests with moving patterns of different spatial wavelengths (S5, S10, S20) at $v = 15.5cm/s$ in the forward and backward direction. Analysis of variances reveals a highly significant influence of the direction of movement ($P < 0.0001$), but no significant influence of the type of pattern ($P = 0.12$). Adapted from [93].

Estimating the shape and distance of objects

In [91], it was also shown that bees can estimate an object's distance by using its apparent motion across the retina when the background is featureless, and probably also its motion relative to the background when the background is textured. When encountering a novel environment comprising unfamiliar objects of unknown size, the bees can evidently use these two types of motion cues to acquire a three-dimensional perception of the visual world.

Other navigation tasks

Biological studies have also shown that the ventral optic flow plays an important role in insects for other navigation tasks, such as altitude control and obstacle avoidance [152, 153] or automatic landing and docking [154] (see [155] for a review).

1.8 Bio-inspired visual sensors

As already said in section 1.6, animals can perform complex tasks in various environments and lighting conditions very efficiently thanks to their visual systems. Although some of these, as, for instance, the fly's visual system, are relatively simple in terms of computational resources, they make it possible to (i) auto-adapt to light while being sensitive to small contrasts and (ii) detect local motion patterns in an very

efficient way. For this reason, many efforts have been made to construct visual sensors that reproduce these bio-inspired models both from a hardware and software point of view.

1.8.1 Wide-Dynamic-Range sensors

A large variety of Wide-Dynamic-Range (WDR) image sensors based on CCD and CMOS technologies has been proposed throughout the years [156], trying to widen the luminosity range as much as the visible spectrum while keeping sensitivity to small changes for every average luminosity in the operating range.

WDR sensors can divided into 7 main categories [156]: companding sensors, multimode sensors, clipping sensors, frequency-based sensors, time-to-saturation (time-to-first spike) sensors, global-control-over-the-integration-time sensors, and autonomous-control-over-the-integration-time sensors.

Companding sensors compress their response to light usually by a logarithmic transfer function [157, 158]. This approach has advantages such as an increased dynamic range, a reduced susceptibility to blooming (an apparent increase in the size of a bright object under overload conditions), and fewer ADC bits for the digital output. In addition, as the logarithmic sensors work in continuous, i.e. non-integrating, mode, they can have a higher time sampling resolution.

Multimode sensors combine a linear and a logarithmic response at dark and bright illumination levels, respectively, by dynamically switching between these two modes of operation [159, 160]. Although both the saturation of linear pixels and the low sensitivity of logarithmic pixels are improved, the sensitivity usually drops by several orders of magnitude after the pixel switches to logarithmic operation.

Clipping sensors apply a adjustment method on the integrating capacitance of each pixel resulting in an adjustment of the integration period of the light flow [161, 162]. This method improve the blooming effect and the sensitivity at high lumininance levels, but the implementation of the reset of the integrating capacitance requires a very accurate analog circuitry and precise timing control.

In frequency-based sensors light intensity is converted into a pulse frequency [163, 164], whereas in time-to-saturation (TTS) (or time-to-first spike) sensors the image is processed according to the time the pixel was detected as saturated [165, 166]. Both types of sensors are based on relatively complex pixels, including ADCs, self-reset circuitry, etc, and therefore the noise floor is usually higher compared to other sensors. In addition, their spatial resolution is usually low since the minimal transistor count within the pixel is relatively higher than other sensors. The power consumption is also high, since the pixel is frequently self-reset and the counter is toggled at the same frequency.

In sensors with global or autonomous control over the integration time, the integration time of each pixel is controlled globally by a predetermined exposure period regardless of its amount of charge or autonomously depending on its amount of charge, respectively [167, 168]. These sensors can achieve the same (and even higher) dynamic range than the previous sensors with higher sensitivity, but they involve more complex processing, such as multiple analog-to-digital conversions and frequent readout cycles from memory units that store the digitized pixel values from previous captures, and are therefore more expensive and power consuming.

Although some of these WDR image sensors can capture images in a luminosity range of up to 7 decades (140 dB), they all provide different contrast sensitivities at different average luminosities. Vision applications involving, for instance, motion detection often require high and constant sensitivity in a large luminosity range, in order to detect small temporal and/or spatial changes in the intensity in several lighting conditions [169, 170]. One possible solution to this problem can be found by looking at the auto-adaptive response of human and animal photoreceptors, as described in section 1.6.1.

Auto-adaptive silicon retinas

Sensors comprising auto-adaptive pixels are often included in the category of companding sensors as their transient response when adapted to light is mostly logarithmic. However, the auto-adaptation process give these sensors a contrast sensitivity which is very different to the WDR sensors presented above, making these sensors belong to a separated category.

The first example of an auto-adaptive silicon retina was presented in [171], where a logarithmic photoreceptor was used to handle transient changes in light in a 1-decade range, while light adaptation within a 1-decade range was obtained by "subtracting" a local spatio-temporal average. This circuit improved the contrast resolution of equally illuminated areas in comparison with standard logarithmic photoreceptor retinas, but there was no improvement in the low signal-to-noise ratio inherent to the logarithmic amplification. In [172], a modified version of this chip was compared with the OPL response described in *Necturus* [173] (see Fig. 1.8), showing light adaptation in a 5-decade range but sensitivity to luminous changes within a range of only 0.5 decades.

To overcome these limitations, a more biologically inspired solution was subsequently developed in a study by [174], which consisted in modulating the synaptic strengths locally to control the sensitivity and including cone-to-cone gap junctions to attenuate the noise (Fig. 1.21(a),(b)). Although the sensitivity was improved in this way from 0.5 to 2 decades, the adaptation to light was not satisfactory because of the circuit deviations resulting from the increasing inter-receptor coupling strength.

A good compromise between contrast sensitivity and light adaptation was reached in [175] (Fig. 1.21(c),(d)), which gave light adaptation in a 6-decade range and sensitivity in a 1-decade range. However, the steady-state response of this pixel was found to increase with the light intensity (i.e., the photodiode current) and the transient response was not always monotonic when large lighting variations occurred (see Fig. 2.13 in [176]). The Delbrück adaptive pixel was also found in studies on optic flow measurements to have little practical use in situations where changes in the light greater than 1 decade are liable to occur (see Figs. 7(b),7(d),7(j) and 7(l) from [177]).

In [178], the Gamma correction method presented in [179] for local tone-mapping purposes was improved by digitally normalizing the pixel output directly in VLSI in line with the Michaelis-Menten law, however only a few preliminary results on light adaptation and contrast sensitivity were presented.

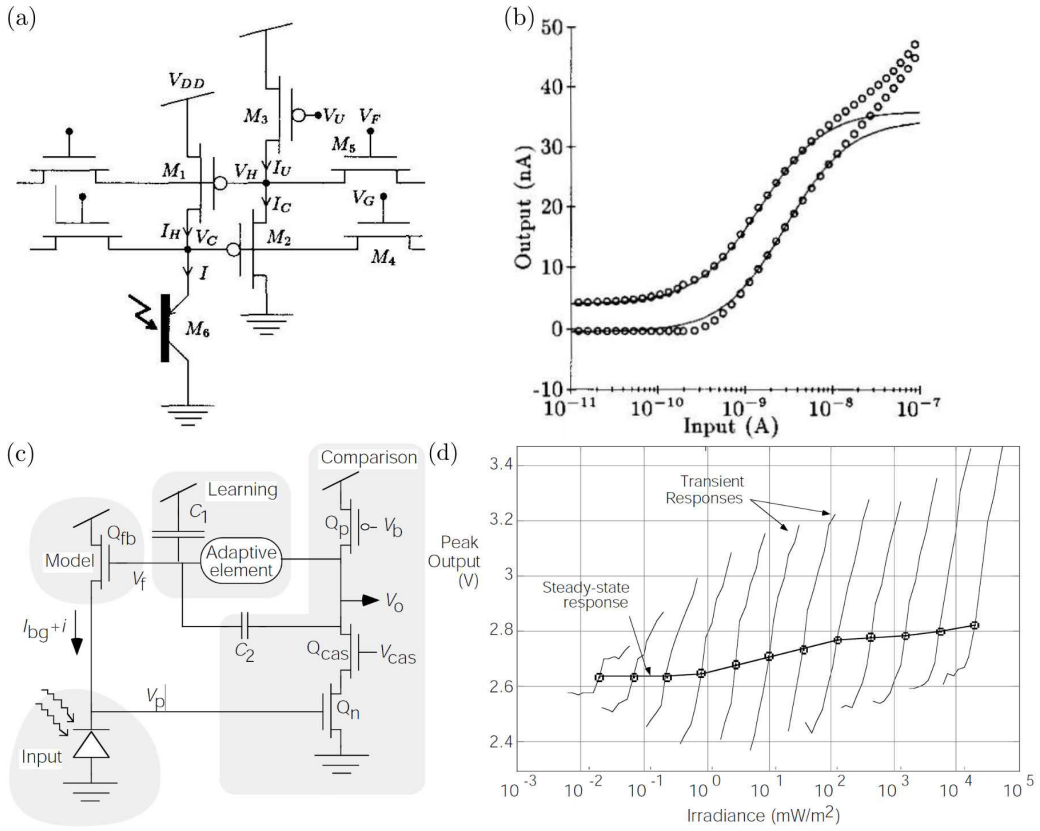

Figure 1.21: (a) The Boahen neuromorphic outer-plexiform pixel circuit and (b) its s-shaped curves characterized at two background intensity levels (the continuous lines are fits of the Michaelis-Menten equation). (c) The Delbruck auto-adaptive pixel circuit and (d) its s-shaped curves characterized at several background intensity levels within a 6-decade range. Adapted from [174] and [175].

1.8.2 Asynchronous event-based sensors

Although active-pixel sensors (APS) have advantages such as having small pixels and compatibility with standard computer-vision algorithms, they also have some clear drawbacks: the pixels are sampled repetitively even if their values are unchanged, the dynamic range of each pixel is limited by its gain and integration time, and short-latency vision problems require high frame rate and produce massive output data. Indeed, the high computational cost and low frame rate of conventional image processing is largely the practical reason for the recent development of neuromorphic address-event representation (AER) silicon retinas. The asynchronous nature of AER sensors output is inspired by the way in which neurons communicate over a large range. Rather than sampling pixel values, AER pixels asynchronously output address events when they detect a "significant" signal (see [180] for a complete review on neuromorphic event-based sensors).

The asynchronous Parvo-Magno retina model presented in 2004 by Zaghloul and Boahen was the first silicon retina modeling both the outer and inner plexiform layers in the retina, including sustained (Parvo) and transient (Magno) types of cells [181].

The spatial contrast and orientation vision sensor, called VISe, presented in 2003 by Ruedi et al. [182] included an on-chip analog computation of contrast magnitude and direction of image features. A temporal ordering of this information according to the contrast magnitude is used to reduce the amount of data delivered. This sensor, realized in a 0.5-μm two-poly three-metal technology, featured a contrast sensitivity of 2%, a contrast direction precision of 3^o, and a dynamic range of 120 dB.

The so-called "Octopus" retina presented by Culurciello et al. in 2003 is an event-based intensity-coding imager that communicates the pixel intensity through the frequency or inter-spike interval of the AER event outputs [164]. The name Octopus comes from the animal because the octopus eye, like other cephalopods, is inverted compared to mammalian eyes, in the sens that the phototreceptors are not behind the optic nerves fibers (as opposed to mammalian eyes), avoiding the presence of a blind spot where the fibers pass through the retina.

In the Dynamic Vision Sensor (DVS) presented in 2008 by Lichtsteiner et al. [183] each pixel quantizes local relative intensity changes independently and in continuous time to generate spike events. These events appear at the sensor's output as an asynchronous stream of digital pixel addresses, giving local intensity changes in the visual scene with sub-millisecond timing precision. The output data rate depends on the dynamic content of the scene and can be orders of magnitude lower or higher than those of conventional frame-based imagers. By combining an active continuous-time front-end logarithmic photoreceptor with a self-timed switched-capacitor differencing circuit, the sensor achieves an array mismatch of 2.1% in relative intensity event threshold and a pixel bandwidth of 3 kHz under 1 kLux. Dynamic range is 120 dB and the chip power consumption is 23 mW. By providing high pixel bandwidth, wide dynamic range, and precisely timed sparse digital output, this silicon retina provides an attractive combination of characteristics for low-latency dynamic vision under uncontrolled illumination with low post-processing requirements (Fig. 1.22(a),(c)).

The Asynchronous Time-based Image Sensor (ATIS) presented in 2011 by Posch et al. combines a DVS pixel with an intensity measurement unit in each pixel in order to trigger light intensity measurements by time-based DVS events [184] (Fig. 1.22(b)). To do this, the 304x240 $30\mu m$-pixel sensor combines the notions of temporal contrast detection [183] and pulse-width modulation (PWM) intensity encoding by using a new time-based correlation double sampling circuit to output the pixel gray level values only from pixels that change in intensity, resulting in a dynamic range of about 130 dB. Neuromorphic pixels of this kind has been found to give very impressive results for surveillance [185] and retinal-implant [186] applications, and are actually used in visual sensors available on the market, such as those provided by Pixium Vision and Chronocam (Fig. 1.22(e)).

However, none of these event-based silicon retinas could provide both spatial and temporal processing in a form suitable for computer vision applications. The new Dynamic and Active-pixel Vision Sensor (DAVIS), which combines DVS and APS technologies in the same pixel [187, 188], seems to be very promising, offering the advantages of both conventional machine vision based on tiny APS and event-based vision based on low-latency, sparse-output neuromorphic silicon retinas (Fig. 1.22(d)).

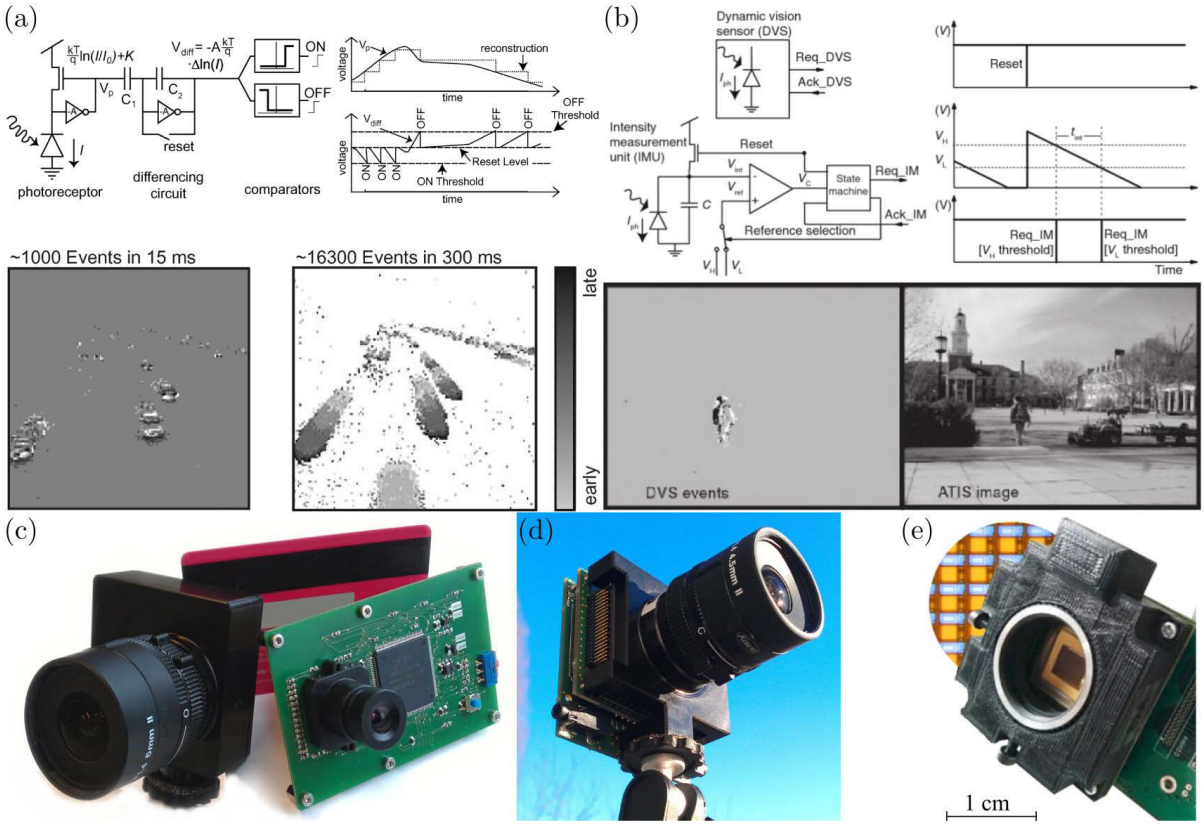

Figure 1.22: (a),(b) Circuit and time response of one pixel of (a) the DVS retina and (b) the ATIS retina, with two examples of images captured by these sensors. Adapted from [183] and [184]. (c) The DVS sensor and (d) the DAVIS sensor developed at the iniLabs laboratory at the ETH University. (e) The Chronocam sensor based on the ATIS retina.

1.8.3 Optic flow sensors

As presented in section 1.7, a huge number of methods have been proposed to compute the global or local optic flow, showing different pros and cons. Since the early 1980s, optic flow have usually been computed by applying these algorithms on image sequences from standard cameras. However, this way of computing the optic flow has 2 main drawbacks: (i) the estimates may be not reliable due to the undersampling or the high dynamic range of the visual scene (see section 1.7.2), and (ii) it often requires very large computational resources to deliver robust measurements. For these reasons, authors have recently developed alternative solutions for measuring the optic flow suitable for real-time applications by adapting on-the-shelf visual sensors with additional hardware or software modules, or constructing new custom-made sensors.

Here, the different technologies which have been used for measuring the optic flow will be presented by dividing them in 3 main categories: CMOS-based sensors, optical-mouse sensors and custom-made sensors based on analog Very Large Scale Integration (aVLSI) chips. The bio-inspired local motion sensors (LMSs) designed and developed at our laboratory will be presented separately in the next subsection.

CMOS-based sensors

One of the first very high-frame-rate optical-flow sensors was developed by Diaz et al. in 2008 using a CMOS camera followed by a superpipelined, fully parallelized architecture for the optical flow processing implemented onboard a FPGA [189]. The 2-D optic flow was estimated by applying a gradient-based method based on the Lucas-Kanade algorithm on images acquired at up to 170 frames/s with a resolution of 800x600 pixels by an onboard DSP. A few years later, another optic flow sensor based on a 320x240 pixels CMOS camera and an onboard FPGA was presented using, this time, a correlation-based algorithm running at 120 fps [190].

In [191], nearly 60000 ERDs were successfully implemented on a modified Eneo SmartCam module, composed of a CMOS camera, a fisheye lens and a FPGA, covering a field of view of 185^o and delivering measurements at 300 fps (Fig. 1.23(a)). However, as discussed in section 1.6.4, the outputs of the ERDs do not provide metric optic flow values and therefore this sensor could be used only as visual ego-rotation sensor for micro aerial vehicles (MAVs).

In 2013, a new optic flow sensor called PX4FLOW, based on a CMOS imager and a block-matching algorithm implemented onboard a ARM Cortex microcontroller, was presented [192], and is now available for purchasing (Fig. 1.23(b)). This compact sensor is also equipped with a rate gyro and an ultrasonic distance sensor to automatically compensate for rotations of the camera and scale the optic flow to metric values, however it delivers only one optic flow measurement at the time instead of a vector field.

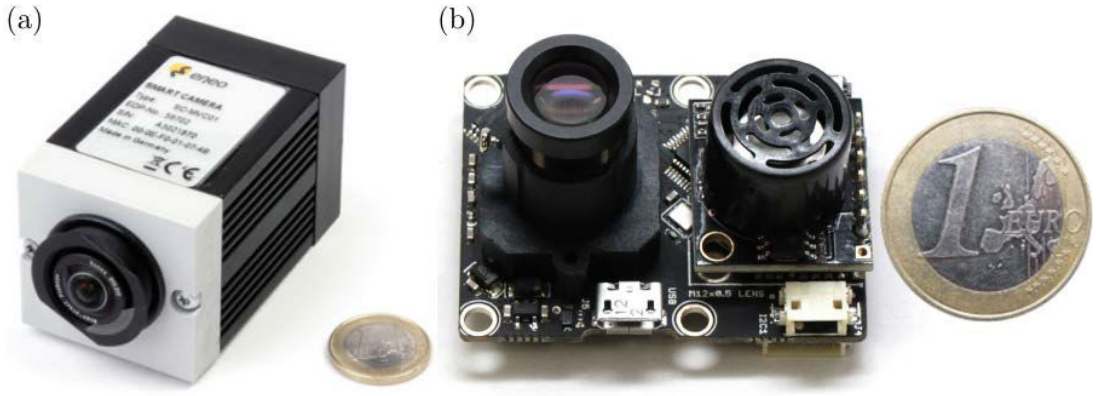

Figure 1.23: (a) The Eneo SmartCam including a CMOS camera, a 185^o fisheye lens and a FPGA, used for implementing nearly 60000 ERDs. Adapted from [191]. (b) The PX4FLOW optic flow sensor with mounted lens on the left and ultra sonic distance sensor on the right. Adapted from [192].

Optical-mouse sensors

The optical mouse presented by Lyon and Haeberli in 1982 [193] was actually the first digital motion detection sensor. In their first chip, the photocurrent of a photodiode was integrated over time on a capacitor and transduced to a digital signal by a simple inverter. The time for charging up the capacitor depended on the input light level. Through lateral inhibition, a cell exposed to a brighter pattern which

reached a high state earlier could suppress the neighbors. Thus, by tracking the location of the winner cells, the movement of the input image could be determined.

Since then, a lot of improvements have been made to the retina and the algorithm used in optical-mouse sensors to process the 2-D local optic flow. Recently, some authors characterized optical-mouse sensors providing analytic models and comparative results [194, 195], however very little is still known about the onboard processing of these sensors. In [196], the authors presented an interesting comparative analysis for tracking a vehicle's position when using optical-mouse sensors instead of standard technologies, such as GPSs, accelerometers or laser rangefinders.

Optical-mouse sensors have been recently used as optic flow sensors onboard wheeled and flying robots by simply adjusting their lens, however they still have some important limitations such as their sensitivity to dynamic lighting conditions and to their distance to the objects (see section 1.9).

Analog VLSI sensors

One of the first analog Very Large Scale Integration (VLSI) sensor for measuring 2-D optic flow was developed by Tanner and Mead in 1986 [197]. This chip was designed to solve the continuity equation in (1.3) for few neighboring pixels, however device mismatches were sometimes produced due to analog division that amplified the noise. A modified version of the Tanner-Mead sensor was presented a few years later in [198], where the temporal derivatives were multiplied instead of divided by the spatial derivatives, in order to reduce the mismatches due to the noise amplification. Later, Stocker presented a different gradient-based implementation on an analog VLSI chip using two-layer networks of locally connected motion units to collectively estimate the optimal 2-D optical flow field [199, 200]. By globally adjusting some parameters, for instance, for varying the smoothness strength, the sensor could provide good continuous-motion estimates.

In 1993, Delbruck presented the first effective implementation of multiple ERDs over a wide spatial and temporal range on an analog VLSI chip, using unidirectional delay lines as tuned filters for moving edges [114]. Other implementations of ERDs on analog VLSI chips were subsequently presented, for instance, in [201–203]. Reviews on conventional gradient-based and correlation-based VLSI chips are given in [204, 205].

In 1995, Kramer et al. [206] presented a 1-D velocity sensor strongly inspired by the time-of-travel scheme presented in [117, 118] (see section 1.7.2). At the time of the occurrence of an edge at one pixel, a capacitor starts charging making its voltage decay logarithmically. When the edge reaches the next pixel, a pulse signal is triggered and the voltage on the capacitor is sampled, delivering an output voltage which is indicative of the edge velocity. A new chip based on this architecture but including a light adaptation process was presented ten years later in [121], where this method was called for the first time "time of travel" (Fig. 1.24). Similar approaches have been presented, for instance, in [207].

Barrows proposed a new design of a motion detection chip based on the Competitive Feature Tracker algorithm, in which the photoreceptors' signals were sent into an array of four feature detectors that

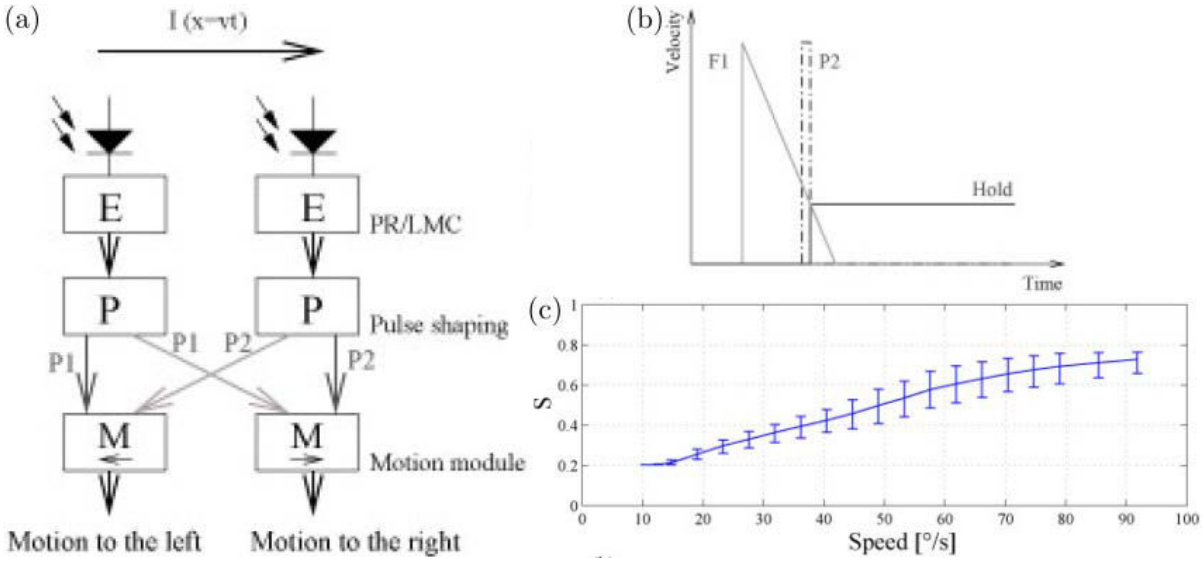

Figure 1.24: (a) Block diagram of the facilitate-and-sample time-of-travel scheme. The E block contains the photoreceptor and laminar monopolar cell (LMC) circuit. Each P block creates a sampling pulse when the LMC output goes between 2 global threshold values. (b) The sampling pulse (P2) resets a capacitor and samples the capacitor voltage (F1) of its neighboring pixel. (c) The sensor's output (S) obtained is proportional to the angular speed. Adapted from [121].

output four analog feature signals [208]. The feature detector circuit reached its highest output value when the feature which it was tuned for appeared on its input photoreceptors. A differential amplifier was used as feature detector, making the motion detectors sensitive to only one direction.

Recently, new-generation optic flow sensors based on the asynchronous event-based retinas presented above have been developed showing very promising results [209, 210].

1.8.4 Local Motion Sensors

Based on the experiments performed on the fly's visual system as presented in section 1.6.4, Franceschini et al. proposed a first analog implementation of a time-of-travel EMD, called 2-pixel Local Motion Sensor (LMS) [117, 118] (Fig. 1.16(b), 1.25(a)). The computation of the time delay Δt between two adjacent photoreceptors' signals was composed of 4 main steps:

- Spatial sampling and low-pass spatial filtering, which was achieved by defocusing a miniature lens placed in front of the photoreceptors to obtain a Gaussian angular sensitivity, as it occurs in the fly's compound eye (see section 1.6.3).
- Band-pass filtering, which resulted from the combination of a low-pass filter to cut off the high frequencies (noise) and a high-pass filter to enhance the transient signals (motion).
- Thresholding with a hysteresis block to detect a contrast. This threshold value was critical to the performances of the LMS: too low values would give wrong measurements (due to the noise),

whereas too high values would decrease the refresh rate.

- Measuring the time delay Δt between two detected contrasts. In this first implementation, the time delay was measured by sampling a decreasing exponential function generated by the first signal after crossing the threshold value, through an impulse generated by the second signal after subsequently crossing the threshold value. The optic flow was therefore approximated by the function $e^{-\frac{\Delta t}{\tau}}$ in order to reduce the noise present at high optic flow values.

Such a 2-pixel LMS could compute the optic flow only in one preferred direction (e.g. the first photoreceptor must see the contrast before the second one). A bi-directional sensor could be implemented by adding a second LMS in the opposite direction and by taking the maximum value between the two measurements [211].

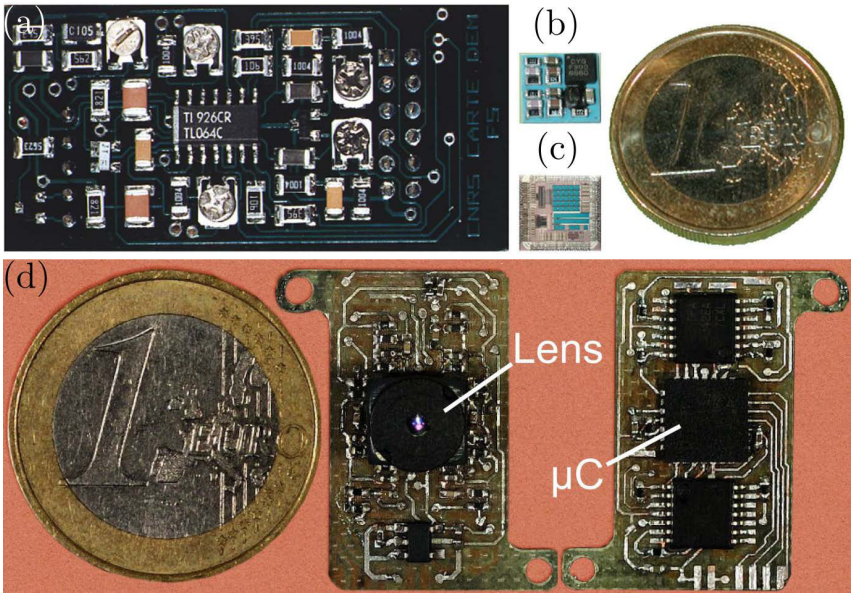

Figure 1.25: Pictures of (a) the first analog version of a LMS developed in 1989 (adapted from [118]); (b) the hybrid analog-digital version of a LMS developed in 2007, involving a tiny microcontroller and constructed using low-temperature cofired ceramic technology (LTCC) (adapted from [212]); (c) the FPGA-based version of a LMS developed in 2007, comprising 25 adaptive photoreceptors (adapted from [213]); (d) the bi-directional 6-pixel LMS developed in 2012 (adapted from [214]).

A few years later, hybrid versions mixing analog and digital processing were proposed [122, 135, 212] (Fig. 1.25(b)). The bandpass-filtering step was still done using analog filtering to differentiate the visual signals but also to avoid aliasing before digitization. The other steps of the processing were carried onboard a microcontroller where the signals were first low-pass filtered to further cut off the high frequencies, in particular the $100\,Hz$ of the artificial lighting, and then two separate pathways were used to process ON and OFF contrasts. The optic flow was finally estimated using an exponential function or a look-up table (Fig. 1.26(a)). This implementation has also been implemented successfully onboard a FPGA [213] (Fig. 1.25(c)).

More recent versions of 6-pixel LMSs, also called Visual Motion Sensors (VMSs), have been developed to increase the range, the precision and the refresh rate of the OF measurements, as well as the luminosity range [177, 214–216]. In [177], the authors implemented the TOT algorithm on both a custom-made VLSI array called APIS (which stands for Adaptive Pixels for Insect-based Sensors) equipped with Delbruck-type auto-adaptive pixels [175], and an on-the-shelf linearly-amplified photosensor/retina called LSC (purchased from iC-Haus). The LMSs were tested for the first time in an outdoor environment, showing that the LSC-based LMS was more accurate in a narrow 1.5-decade illuminance range, whereas the APIS-based LMS was more robust to illuminance changes in a larger 3-decade range. In [214], a bi-directional version of the LSC-based LMS was presented by applying the same maximum criterion as presented in [211] (Fig. 1.25(d), Fig. 1.26(b)). In [216], a modified version of the LSC-based LMS was used outdoors on an helicopter showing that it was possible to adapt the OF range of a LMS, in particular to measure low optic flow values, by adjusting the optical lens and the filtering steps.

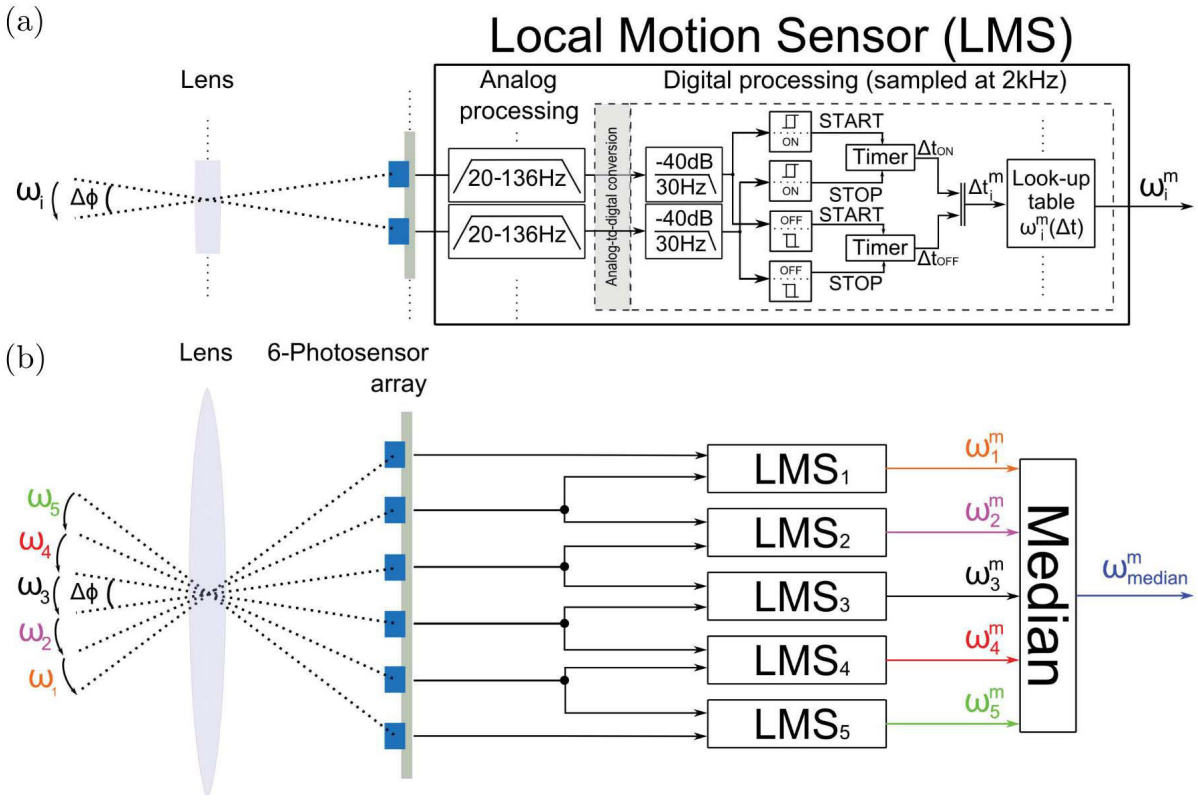

Figure 1.26: (a) Scheme of a hybrid analog-digital 2-pixel LMS. Steps 1 and 2 were performed prior to digitization. The output from each analog filter was digitized through an analog-to-digital converter (ADC) and the other steps of the processing were realized onboard the microcontroller. Two separate pathways process the angular speed of ON and OFF contrasts. (b) Scheme of a bi-directional 6-pixel LMS. First, the median value of the 5 optic flow values measured between each pair of adjacent photoreceptors was taken for each direction, then the maximum values between the 2 median values was taken as the final output. Adapted from [214].

Although they have shown very good results in certain testing conditions, the threshold-based TOT methods used in these works have some important drawbacks. First, as explained above, the performances of the LMS strongly depend on the thresholding phase, which makes the sensor dependent of the operating conditions, such as the visual patterns, the light levels and possible vibrations. Then, the refresh rate of the OF measurements depends not only on these operating conditions but also on the bandwidth of the band-pass filter, whose high cut-off frequency is relatively low to cut off the noise.

A different implementation of a 2-pixel LMS has been proposed by continuously tracking the time delay Δt between two neighboring signals using a gradient-based approach similar to the Lucas-Kanade method [217]. Although not bio-inspired, this processing allowed to process the angular speed in both directions with good accuracy without being dependent on a threshold value. However, it was more costly in terms of computational power and was sensitive to the initial condition on the time delay, which influences the convergence of the processing.

Although showing very promising results, the LMSs presented so far require some important improvements in order to be used outdoors on a real vehicle in unknown, unstructured environments under dynamic lighting condition within a wide range.

1.8.5 Artificial compound eyes

As described in section 1.6.3, the compound eye can deliver very effective motion detection over a very wide field of view (FOV) using less pixels than standard cameras, thanks to its optically-isolated facets (e.g. 800 to 900 facets over nearly 360^o in the fly). Therefore, artificial implementation of compound eyes has attracted researchers during the last few years. The main challenge in the the design of an artificial compound eye is to be able to manufacture an array of microlenses that can be precisely aligned with photosensitive layers. Recently, Song et al. (2013) proposed a solution based on an elastomeric microlens array placed on top of a stretchable array of photodiodes interconnected by filamentary serpentine wires deformed into an hemispherical shape [218] (Fig. 1.27(a)). In the same year, Floreano et al. presented another biomimetic compound eye featuring a panoramic, undistorted field of view by using three planar layers of separately produced arrays, namely a microlens array, a neuromorphic photodetector array, and a flexible printed circuit board that are stacked, cut and curved to produce a mechanically flexible imager [219] (Fig. 1.27(b)).

1.9 Bio-inspiration in robotic and automotive applications

Researchers and engineers have tried for long time to take inspiration from nature for developing machines that can help human lives in the most efficient and effective way. Biologically-inspired technologies can be found in a wide range of application fields, from medical to aerospace applications, but all answering the same question: how can nature solve certain problems and how can its solutions be implemented on electromechanical devices.

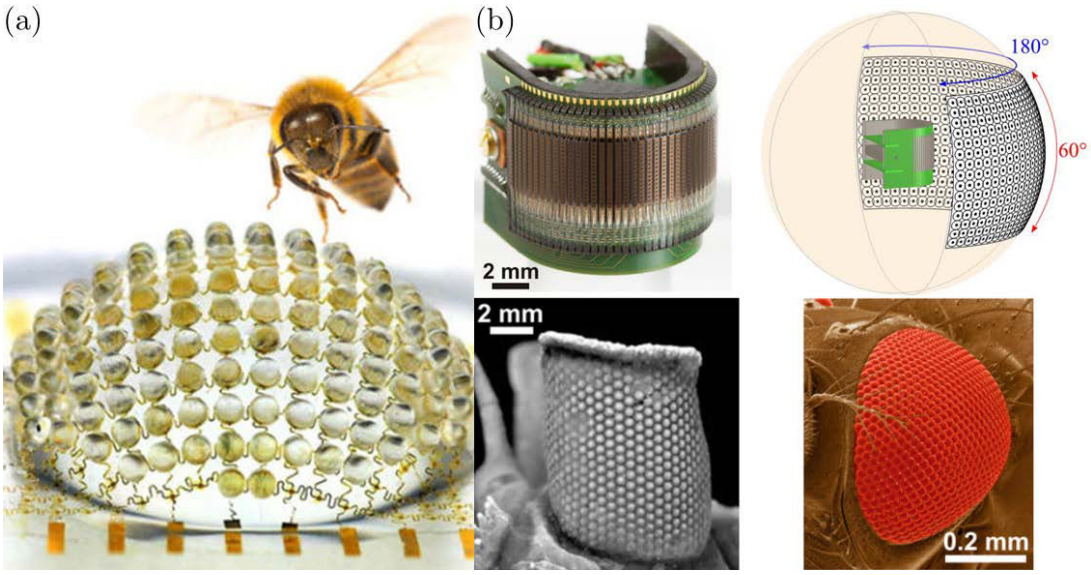

Figure 1.27: (a) The artificial compound eye presented in [218] with a picture of a honeybee. (b) The curved artificial compound eye (CurvACE) presented in [219], with a schematic representation of its FOV and two pictures of the large compound eye of an extinct trilobite and the minute eye of the extant fruitfly Drosophila.

In robotics, there exists two approaches: bio-inspiration and biomimetics. In the former, the aim is to take inspiration from nature to develop new effective technologies by realizing an energy-efficient technical adaptation of the natural model but without necessarily copying nature. Wheeled robots, and eventually vehicles, using bio-inspired sensors (see, for instance, section 1.8) or actuators are good examples of this approach. Although these sensors and actuators are based on biological findings, they are not constructed and employed to mimic the animal's behavior but to improve the performances or reduce the cost of the machine which they are applied to. In the latter, the aim is to go a step further trying to mimic the animal's behavior and, by looking at the robot's response, sometimes trying to better understand the animal itself and therefore reformulate some hypothesis that have been made on the biological models. This is the reason why researchers are now developing robots endowed with biomimetic locomotion systems to walk, crawl, swim or fly based on nature's best solution found through thousand of years of evolution. Although the biomimetic approach is not directly related to this thesis, we would like to give in the following section a short overview on biomimetic robots for completeness.

1.9.1 Biomimetic robots

The BionicANTs presented by Festo in 2015 are a well-representative example of biomimetic robots since they not only reproduce the delicate anatomy of the ants but, for the first time, they also mimic the cooperative behavior of these creatures using complex control algorithms (Fig. 1.28(a)).

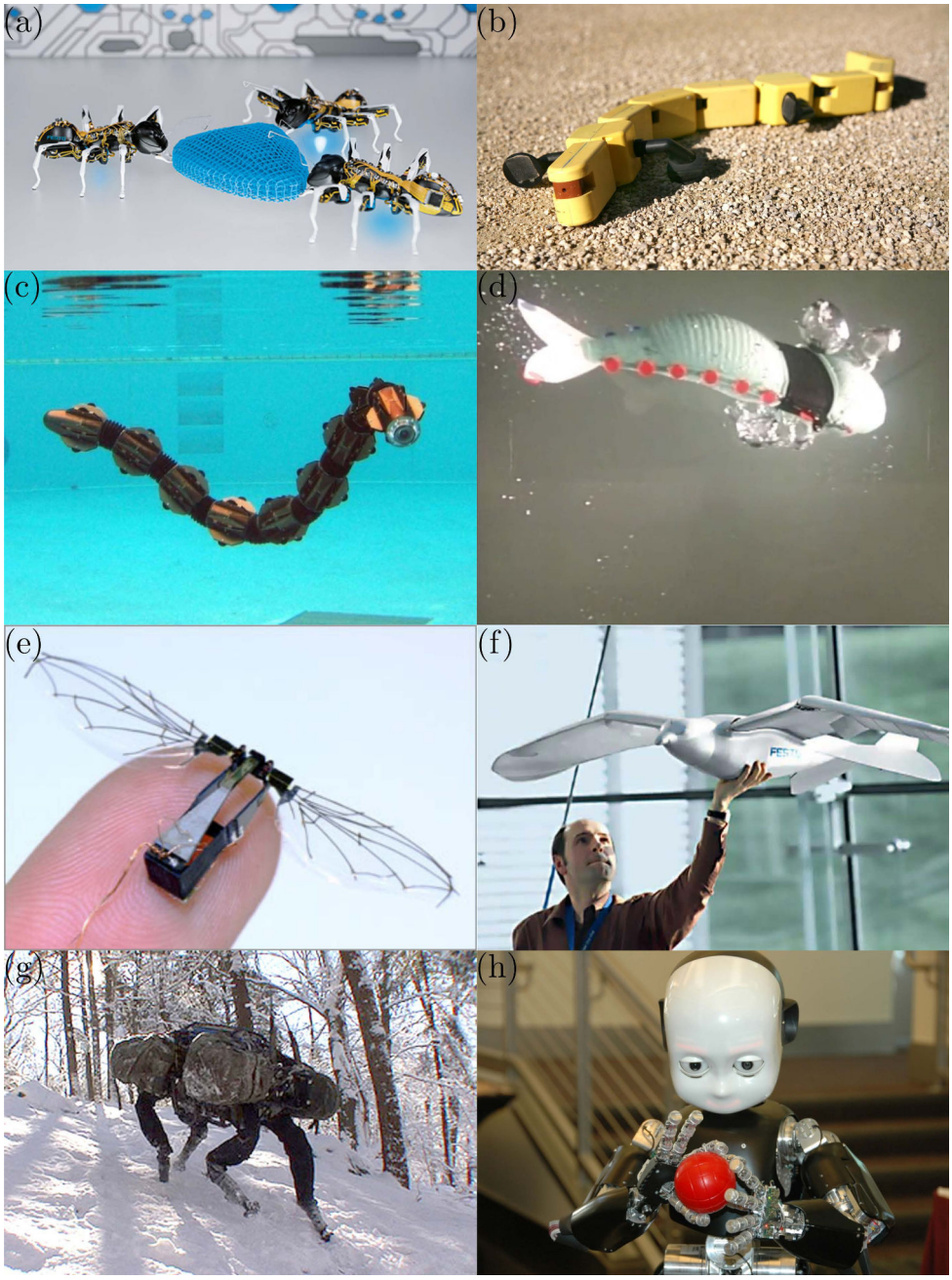

Figure 1.28: (a) The BionicANT robots developed by Festo. (b) The Salamandra Robotica II presented in [220]. (c) The ACM-R5 snake-like robot presented at EXPO2005 by HiBot. (d) The fish-like soft root presented in [221]. (e) The insect-sized RoboBee robot presented in [222]. (f) The SmartBird robot developed by Festo. (g) The BigDog quadruped robot developed by Boston Dynamics. (h) The iCub humanoid robot developed by the Italian Institute of Technology.

As the transition from aquatic to terrestrial locomotion was a key step in vertebrate evolution, various amphibious robots that mimic the salamander and snake locomotion have been presented, for instance, in [220,223,224] to bring light on their ability to switch between swimming and walking (Fig. 1.28(b),(c)).

Another example of biomimetic aquatic robot is the soft-bodied robot developed by Marchese et al. (2014), which is capable of rapid, continuum-body motion by employing a compliant body with embedded actuators emulating the slender anatomical form of a fish [221] (Fig. 1.28(d)).

By looking at the flapping-wing aerodynamics of insects, Wood et al. (2008) developed the first insect-sized (60mg) Micro Air Vehicle (MAV) capable of Diptera-like wing trajectories by producing sufficient thrust to accelerate vertically [222] (Fig. 1.28(e)). They successfully embedded on such an ultra light MAV an optic flow sensor in order to estimate its altitude, making it able to hover and perform some basic controlled flight maneuvers [225]. On similar lines, the SmartBird robot developed by Festo in 2011 is another ultralight but powerful flying robot with excellent aerodynamic qualities and extreme agility by mimicking the flight of birds (Fig. 1.28(f)).

Various quadruped robots that move and run like dogs, as well as anthropomorphic robots that walk, manipulate and learn like humans have also been presented during the last decade. Probably some of the best examples of these robots are those developed by Boston Dynamics, as they provide extremely stable and natural gates in extreme terrain conditions, such as walking on very steep and slippery surfaces including snow, ice and mug. In particular, the BigDog robot, founded by DARPA and presented in 2010, was the first quadruped robot that could walk, run, climb and carry heavy loads on rough terrains by reproducing the dog's locomotion using compliant elements to absorb shock and recycle energy from one step to the next (Fig. 1.28(g)).

The iCub robot developed by the Italian Institute of Technology is the first humanoid robot that can learn how to lift up when lying on a floor and move on 4 legs, or how to manipulate objects and interacting with the environment in a very child-like manner (Fig. 1.28(h)).

1.9.2 Bio-inspired wheeled robots

One of the first bio-inspired optic flow-based robots was developed by Srinivasan et al. (1999), based on the behavioral results obtained previously on honeybees (Fig. 1.29(a)). The small wheeled robot was equipped with a forward-facing camera which captured views of both side walls (left and right) using two mirrors in front of it [115]. The optic flow produced by the relative motion of the robot with respect to the two side walls was first computed on a desktop computer using a simplified version of the image-interpolation algorithm presented in [111] (see section 1.7). Then, suitable steering commands were delivered to the robot to keep its trajectory close to the midline of the tunnel, reproducing a similar behavior to that observed in their experiments on honeybees [91, 92].

These authors also implemented on the same robot a localization method inspired by the bee's visual odometer [146], to determine the distance traveled by the robot by integrating over time the optic flow experienced in the two sides. This computation did not yield the absolute travel distance since this

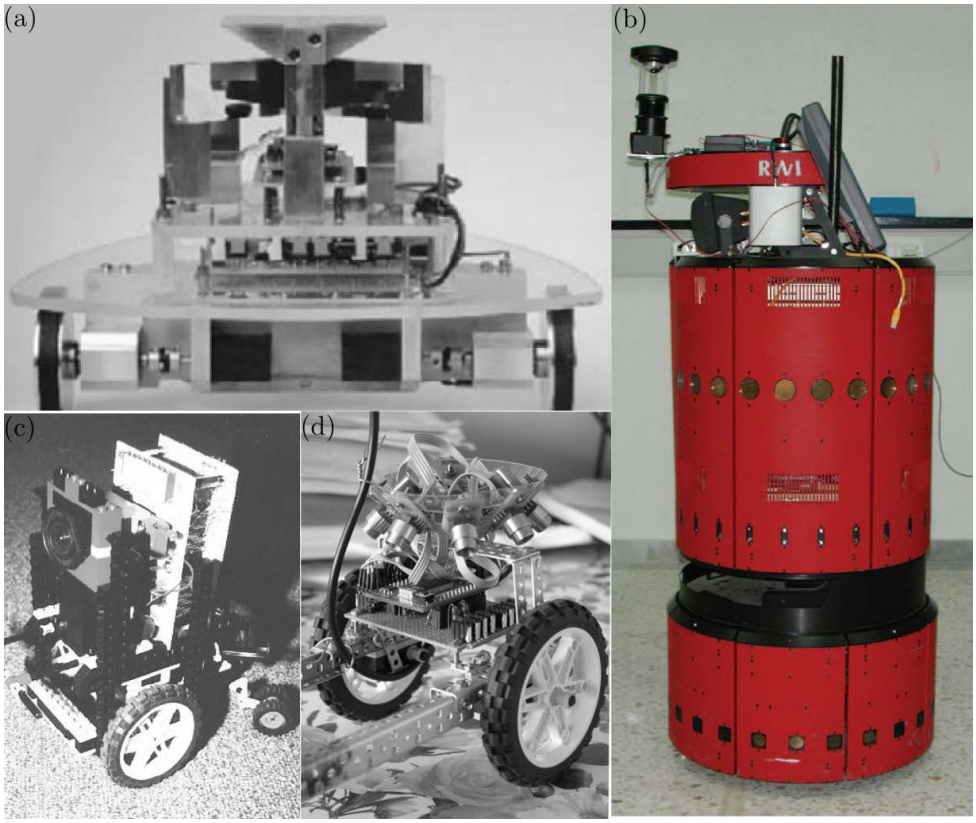

Figure 1.29: Examples of bio-inspired wheeled robots. (a) Corridor-following robot presented in [115], which was able of traveling in a tunnel by balancing the optic flow measured from the two sides. (b) The mobile robot presented in [226] with a panoramic camera mounted on it, which was able of navigating in corridors by either staying in the middle or following one wall. (c) The asymmetric wheeled robot presented in [227] with an aVLSI motion sensor based on ERDs, which could compensate its yaw rotations and maintain straight trajectories thanks to the optomotor feedback. (d) The wheeled robot presented in [228] with a head containing eight optical mouse sensors looking down to the ground at about 45^o to estimate its trajectory by integrating the optic flow measured.

depends on the distance to the walls (unknown), but it still yielded a value indicative of distance traveled [115]. To overcome this shortcoming, the authors designed another robot which, this time, did not compute image motion per se. Instead, a modified version of the image-interpolation algorithm was applied to two panoramic views of the world captured simultaneously from a fixed, known distance apart on the robot's translating axis [115].

To decrease the important computational power required to extract visual information, [227] created a small, low-power visual sensor with integrated analog parallel processing to extract motion in real-time based on ERDs. This sensor was used on a mechanically asymmetric wheeled robot compensating its yaw rotations thanks to a control feedback loop mimicking the optomotor response of insects (Fig. 1.29(c)).

Similar approaches but using different sensors for computing the optic flow or different estimation and control algorithms have been presented, for instance, in [226, 229, 230]. In [226], authors pointed

out that such strategy would lead to a crash of the robot on featureless walls and therefore they implemented a wall following behavior triggered by the absence of optic flow measurements from one side (Fig. 1.29(b)).

Humbert et al. proposed a new solution to estimate the position and velocity of a vehicle by integrating over a wide field of view local optic flow measurements using their Fourier-series decomposition [231]. This method, called Wide-Field Integration (WFI), was inspired by the LPTC cells of the fly (see section 1.6.3) and was successfully tested on a wheeled robot traveling inside a corridor to center the robot and control its forward speed [232]. The method was later extended to a 6-DOF robot to navigate in unknown environments [233] and implemented onboard a quadrocopter [234]. However, in all these experiments, the optic flow measurements were coupled with other sensors, such as a sonar, an accelerometer or a speed sensor, to obtain metric measurements.

Dahmen et al. (2009) presented a cheap and light wheeled robot that extracts self-motion exclusively from optic flow, inspired by both the eye structure and the visually-controlled behavior of water striders [228]. Optical mouse chips provided with adequate lenses were mounted on the robot's head looking at the ground over 360^o, and used for measuring the optic flow induced by the visual motion of the ground (Fig. 1.29(d)). Experimental results of the odometry method proposed were first provided by manually driving the robot indoors along a straight line and around its yaw axis. Then, a different configuration of the sensors was tested outdoors on a moving arm instead of the robot.

1.9.3 Bio-inspired robots developed at our laboratory

Based on the physiological and behavioral studies carried out on the fly and particularly on its visual system (see section 1.6.4), Franceschini et al. built in 1989 the first bio-inspired vision-based wheeled robot called "robot mouche" (fly robot). The fly robot was composed of 118 analog electronic circuits modeled upon the neural circuitry of the fly's brain connected to a circular array of photoreceptors and was capable of approaching a goal while avoiding obstacles on its way by reproducing a behavior similar to the fly's optomotor response [118, 119] (Fig. 1.30(a)). In particular, the robot could travel in translation at a known forward speed ($50\,cm/s$) while controlling its direction and determining its distance to the objects by using only 114 EMDs.

Ten years later, a similar control strategy was implemented on a 850g flying robot called FANIA which was able to fly away from obstacles and follow a terrain without any knowledge of its speed but using only 20 local optic flow measurements [235, 236].

Inspired by the findings obtained on insects (see section 1.7.3), Ruffier et al. proposed in 2003 an autopilot called OCTAVE, which stands for Optical altitude Control sysTem for Autonomous VEhicles, to control the altitude of a flying robot by using only a 2-pixel optic flow sensor placed on the ventral part of the robot facing downwards [122, 237–240]. This autopilot was successfully implemented on a 100g tethered rotorcraft (Fig. 1.30(b)) flying over a contrasted ground and performing challenging tasks such as taking-off, landing, terrain following and reacting suitably to wind.

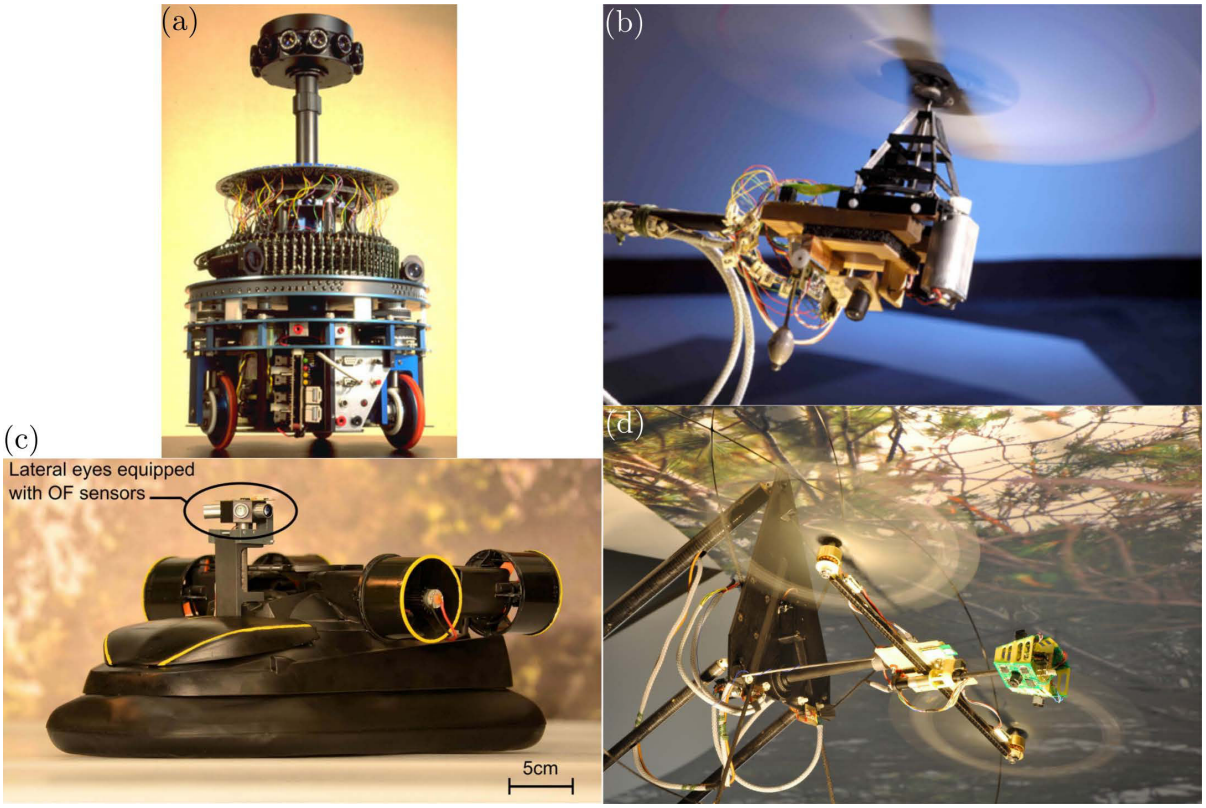

Figure 1.30: (a) The "robot mouche" (fly robot). Reprinted from [60]. (b) The OCTAVE robot. Copyright H. Raguet. (c) The LORA robot. Reprinted from [241]. (d) The BeeRotor robot. Reprinted from [242].

Based on their results from behavioral experiments on honeybees (see section 1.7.3), Serres et al. (2008) proposed a bio-inspired autopilot called LORA III, which stands for Lateral Optic Flow Regulation Autopilot III, consisting of a dual OF regulator to control both forward and lateral speed by using 2 separate OF set-points [135]. The first lateral OF regulator adjusts the robot's lateral speed so as to keep the lateral OF equal to the sideways OF set-point. Therefore, the distance to the wall becomes proportional to the robot's forward speed, depending on the sideways OF set-point: the faster it travels, the further away from the walls it will be. The second lateral OF regulator adjusts the robot's forward speed so as to maintain the sum of the two (right and left) OFs equal to the forward OF set-point. Thanks to these two intertwined feedback loops, the vehicle forward speed is automatically adjusted to the size of the corridor, depending on the forward OF set-point: the narrower the corridor is, the slower it travels. This autopilot has been first tested in simulation [135] and then implemented on a 800g miniature hovercraft traveling in a corridor [241] (Fig. 1.30(c)).

To generalize the OCTAVE and LORA autopilots in order to control the position and speed of an agent freely-flying in a 3-D environment, the ALIS autopilot, which stands for AutopiLot using an Insect based vision System, has been subsequently proposed and tested with a simulated flying agent [243].

Lastly, Expert et al. recently developed a bio-inspired tandem rotorcraft called BeeRotor mimicking optic flow-based behaviors and flying in the vertical plane in a highroofed tunnel depicting natural

scenes [242] (Fig. 1.30(d)). The 80g tethered miniature robot was able to control its heave, surge and pitch, while reorienting its quasi-panoramic visual system made of only 24 pixels to be perpendicular to the followed surface. The BeeRotor robot was able to perform complex tasks such as ground and ceiling hugging and speed control in the presence of a rugged or unstationary environment without any accelerometer or estimation of its pitch, altitude or groundspeed. Its robustness was shown even in absence of the airspeed sensor or when strong perturbations were applied on the aircraft.

1.9.4 Bio-inspiration in automotive applications

Although research works on autonomous vehicles have been presented since the early 1980s (see section 1.5), only recently bio-inspired technologies have been proposed for ADAS systems.

Sota et al. (2004) proposed a motion detection system for a blind spot overtaking monitor which was based on Reichardt correlators [244] (see section 1.6.4). The approach presented used the saliency of motion features in a competition scheme to filter out noise patterns and make features corresponding to rigid body motion self-emerge from the background. The system was tested with real overtaking image sequences, but no real-time implementation on a vehicle was presented.

In [245], the authors presented an architecture and retino-topic unit based on a bio-inspired high-dynamic-range (HDR) sensor intended for automotive applications, in which collision threats were identified by mimicking the neuronal structure of the Lobula Giant Movement Detector (LGMD) found in the locust's eye. This chip and collision detection system were subsequently integrated into a commercial car (Volvo XC90) and tested to deliver collision warnings in real traffic scenarios [246].

In the framework of the European project ADOSE, which stands for Application-specific Detection of road users with vehicle On-board SEnsors, Kogler et al. (2009) presented a stereo vision system for use in pre-crash warning applications for side impacts based on an event-based silicon retina inspired by the human visual system [247]. A block-scheme stereo vision algorithm providing distance information extracted from moving objects was successfully implemented on a DSP satisfying real-time constraints, however no results on a real vehicle usage was presented.

Michalke et al. (2009) presented a biologically-motivated ADAS system based on the generic principle of attention, inspired by front-end visual signal-processing principles in the human brain, to perform an early task-dependent pre-selection of interesting image regions and decrease the scene complexity [248]. The system proposed was subsequently implemented on a Honda Legend prototype using a standard CCD camera and successfully tested on typical construction sites on highways [249].

An ADAS system inspired by the fish schooling behavior has been recently proposed to provide swarm-behavior capabilities to autonomous vehicles using only local knowledge of their environment [250]. This system was first implemented on a dynamic driving simulator to gather driver's feelings and then in traffic simulation software to evaluate gains obtained for a set of equipped vehicles.

1.10 Thesis overview

The aim of this thesis is to: (i) evaluate the potential interest of the local motion sensors (LMSs) on outdoor mobile robots and automobiles, (ii) develop new sensors that are adapted to the road environment, and (iii) propose some applications in which these sensors could improve the ADAS systems of tomorrow.

All but the last chapter feature either published or submitted papers, each dealing with separated but related works. The papers were adjusted to comply with the format of this manuscript, however their contents have not been altered.

In chapter 2, we present how the bio-inspired optic flow sensors developed at our laboratory (LMSs) could be used in automatic parking maneuvers. For this purpose, we developed a low computational-cost method of detecting and tracking a parking spot in real time using 1-D OF measurements around the vehicle together with the vehicle's longitudinal velocity and steering angle. Highly simplified 2-D parking simulations were first performed using Matlab/Simulink software, then some preliminary experiments were carried out using a vehicle equipped with two 6-pixel LMSs.

In chapter 3, we present a novel bio-inspired auto-adaptive silicon retina designed and developed by our laboratory in collaboration with the Center of Particle Physics of Marseille (CPPM), that reproduces the auto-adaptation process found on the animal and human retina. We successfully tested this silicon retina, showing that the novel pixel, called M^2APix, which stands for Michaelis-Menten Auto-Adaptive Pixel, can auto-adapt in a 7-decade range and respond appropriately to step changes up to ± 3 decades, while keeping sensitivity to contrasts as low as 2%.

In chapter 4, we present a novel optic flow sensor based on this auto-adaptive retina and a new robust method for computing the optic flow, which provides several advantages to previously developed optic flow sensors such as its robustness to high-dynamic-range lighting conditions, various visual patterns and vibrations. To test the performances of this novel sensor and show how it can be used for robotic and automotive applications such as visual odometry, we constructed a low-cost car-like robot, called BioCarBot, which stands for Bio-inspired Visually-guided Car-like Robot. This robot can estimate robustly its velocity and steering angle using only the optic flow measurements delivered by two downward-facing sensors along with an Extended Kalman Filter (EKF). Indoor and outdoor experiments were successfully carried out in a 7-decade light level range and using various textures, showing promising perspectives of these sensors for odometry-based applications.

In chapter 5, we present some conclusions on the work accomplished and discuss about further improvements that could be carried out on next versions of LMSs or on the methods proposed here for automatic parking and visual odometry. A new multi-facets sensors composed of 4 M^2APix chips on a flexible Printed Circuit Board (PCB) is also presented, along with some perspectives on other applications as well as on the integration of the LMSs on cars.

Chapter 2

Bio-Inspired Optic Flow Sensors for Automatic Parking Maneuvers

2.1 Introduction

As already mentioned in section 1.4, most intelligent parking assist (IPA) systems available on the market are designed to help the driver with a parallel and/or perpendicular parking maneuver by using ultrasonic sensors and rear-view cameras [8–10] or wide-angle cameras all around the vehicle [11, 12].

Although some IPA systems can already perform (semi-)automatic maneuvers, these systems have some clear limitations, mainly due to the limits of the sensors used (see section 1.3), that do not allow safe fully-automated maneuvers. Such limitations could be summarized as a common problem for mobile autonomous systems (see section 1.2): they can not provide both a wide field of view and a high operating frequency satisfying safety and real-time constraints.

On the one hand, the ultrasonic sensors have a narrow field of view (25 to 45^o), a short-distance range ($< 5m$) and low precision (5 to 15%), therefore, even if many of these sensors are used to cover the area surrounding the vehicle, it is not possible to precisely detect and track the parking spot for a fully-automated maneuver. On the other hand, cameras can cover wide field of views (e.g. fish eye cameras), but they suffer from high-dynamic-range lighting conditions and the image processing require too large computational resources to extract all the information required about the environment in real time. In both cases, the refresh rate, i.e. the frequency at which useful information (distance, velocity, etc) is extracted, ranges from 10 to $30\,Hz$ in standard vehicles.

Fully-automatic parking maneuvers have been presented in various research works [13, 15] but they still require too costly sensors, such as high-resolution lidars and radars, as well as large computational resources to reconstruct the surrounding 3-D environment in real time during the maneuvers.

As described in section 1.8.4, our laboratory have been developing for nearly 20 years 1-D optic flow sensors, also called Local Motion Sensors (LMSs), inspired by fly's visual system. These low-cost sensors

can deliver high-frequency measurements using a very few pixels and simple embedded electronics, and have been successfully used for robotic applications where 2-D trajectories where performed in closed loop. However, the LMSs developed so far, based on the time-of-travel scheme presented in sections 1.7.2 and 1.8.4, had some important limitations due to the thresholding phase: too low threshold values would give wrong optic flow measurements due to the noise, whereas too high values would decrease the refresh rate. This means that, due to this threshold value and the band-pass filter (see section 1.8.4), the reliability and the refresh frequency of the optic flow measurements strongly depended on the operating conditions (visual patterns, light levels, vibrations, etc).

Such an issue was confirmed after testing the LMSs presented in [177,216] outdoors on a real vehicle in different environmental conditions. We therefore designed and implemented a new robust time-of-travel algorithm for estimating the optic flow based on a cross-correlation method similar to that proposed in [123] (see section 1.7.2). This algorithm was successfully implemented in RTMaps software on a desktop PC, and was then applied to the photoreceptors' signals acquired from the LMSs presented in [177,216]. The details of this new method, which has been subsequently improved, will not be presented in this chapter but in chapter 4.

In the work presented in the following paper, we investigated how LMSs could be used to improve automatic parking maneuvers, and eventually represent an alternative or complementary solution for new IPA systems. For this purpose, we developed a low computational-cost method of detecting and tracking a parking spot in real time using 1-D OF measurements around the vehicle together with the vehicle's longitudinal velocity and steering angle. Highly simplified 2-D parking simulations were first performed using Matlab/Simulink software, then some preliminary experiments were carried out using a vehicle equipped with two 6-pixel LMSs based on the new time-of-travel algorithm.

Some details on the mathematical modeling and the experimental results presented here will be given in Appendix A.

2.2 Article

"Towards an Automatic Parking System using Bio-Inspired 1-D Optical Flow Sensors"

Article published in *Proceedings of the 2015 IEEE International Conference on Vehicular Electronics and Safety*, Yokohama, Japan, 2015

Authors: Stefano Mafrica,[1, 2] Alain Servel[2] and Franck Ruffier[1]

2.2.1 Abstract

Although several (semi-) automatic parking systems have been presented throughout the years [13, 15, 251–260], car manufacturers are still looking for low-cost sensors providing redundant information about the obstacles around the vehicle, as well as efficient methods of processing this information, in the hope of achieving a very high level of robustness. We therefore investigated how Local Motion Sensors (LMSs) [216, 242], comprising only of a few pixels giving 1-D optical flow (OF) measurements could be used to improve automatic parking maneuvers. For this purpose, we developed a low computational-cost method of detecting and tracking a parking spot in real time using 1-D OF measurements around the vehicle as well as the vehicle's longitudinal velocity and steering angle. The algorithm used was composed of 5 processing steps, which will be described here in detail. In this initial report, we will first present some results obtained in a highly simplified 2-D parking simulation performed using Matlab/Simulink software, before giving some preliminary experimental results obtained with the first step in the algorithm in the case of a vehicle equipped with two 6-pixel LMSs. The results of the closed-loop simulation show that up to a certain noise level, the simulated vehicle detected and tracked the parking-spot assessment in real time. The preliminary experimental results show that the average refresh frequency obtained with the LMSs was about 2-3 times higher than that obtained with standard ultrasonic sensors and cameras, and that these LMSs therefore constitute a promising alternative basis for designing new automatic parking systems.

2.2.2 Introduction

Since the pioneering works by Paromtchik *et al.* [251] and the first commercial Intelligent Parking Assist System (IPAS) [252], a large range of (semi-) automatic parking methods have been developed over the years using all kinds of sensors, from ultrasonic sensors [253–255] to cameras [256–258], laser scanners [13, 259], and radar systems [15, 260].

On the one hand, high-performance automatic parking systems require costly sensors, such as laser scanners [13,259] and radars [15,260] and large computational resources to reconstruct the surrounding

[1] Aix-Marseille Université, CNRS, ISM UMR 7287, 13288 Marseille, France
[2] PSA Peugeot Citroën, Route de Gisy, 78140 Vélizy-Villacoublay, France

3-D environment in real time during parking maneuvers.

On the other hand, most of the (semi-) automatic parking pilots available on the market involve the use of ultrasonic sensors and rear-view cameras [8–10, 255, 257] to detect a free parking spot while the driver is driving along a row of parked vehicles and to autonomously control the steering system by estimating the vehicle's ego-position via odometry.

To improve IPAS systems, recent research works and industrial developments using Around View Monitor Systems (AVMSs) have been proposed [11, 12, 258]. AVMSs mainly use wide-angle cameras (such as fish-eye and catadioptric cameras) directed toward the ground, giving a view of the surrounding environment, to detect and track parking-spots' ground marks during parking maneuvers.

However, the autonomous vehicles of the future will require more accurate redundant information than that provided by the systems currently available on the market, mainly because of (i) the narrow field of view (FOV) and the short distance range of the ultrasonic sensors used and (ii) the low luminosity range and the low sampling frequency of standard cameras.

As far as we know, the challenge for designing the IPASs of future autonomous vehicles consists in finding low-cost, fast-response sensors providing redundant information about the obstacles present around the vehicle, combined with a sensor-data processing system requiring very low computational resources.

We therefore investigated how Local Motion Sensors (LMSs) [216, 242], giving 1-D optical flow (OF) measurements using methods based on findings obtained on the fly's visual system [62] could be used to improve automatic parking maneuvers.

Thanks to their wide luminosity range and their sensitivity to small contrasts [242, 261], LMSs comprising only a few autoadaptive pixels seem to constitute a promising alternative to standard cameras in situations where the lighting conditions are highly variable and the visual patterns created by the vehicle's body, for example, show small contrasts. In addition, an array of LMSs would have the following advantages with respect to ultrasonic sensors and cameras: (i) faster responses (up to 200 Hz); (ii) a custom FOV (such as 180^o); (iii) a longer distance range than ultrasonic sensors; (iv) less computational cost than cameras.

In this paper, we present the first results toward developing a low computational-cost method of performing automatic parking maneuvers whereby a parking spot can be detected and tracked in real time based on a visual motion sensor setup performing 1-D OF measurements around the vehicle. The algorithm used for this purpose was composed of 5 processing steps, as described in Sec. 2.2.5.

As proof of concept, some results obtained in a simplified 2-D parking simulation implemented in Matlab/Simulink in closed loop will first be presented. Then, to report on the work in progress, we will give some preliminary experimental results corresponding to the first step in the algorithm in the case of a vehicle equipped with two 6-pixel LMSs on a vehicle.

In Sec. 2.2.3, the principles underlying an elementary LMS and the visual-sensor setup used in the simulation will be described. In Sec. 2.2.4, the point and line motion model used will be presented. In

Sec.2.2.5, the various steps in the algorithm we developed will be outlined. In Sec. 2.2.6 and 2.2.7, we will present and discuss the results of some simulations and preliminary outdoor experiments. Some conclusions will be reached in the last section.

2.2.3 Local motion sensors for parking maneuvers

Figure 2.1 shows the principles underlying a 2-pixel LMS [122] and a photo of a 6-pixel LMS [216].

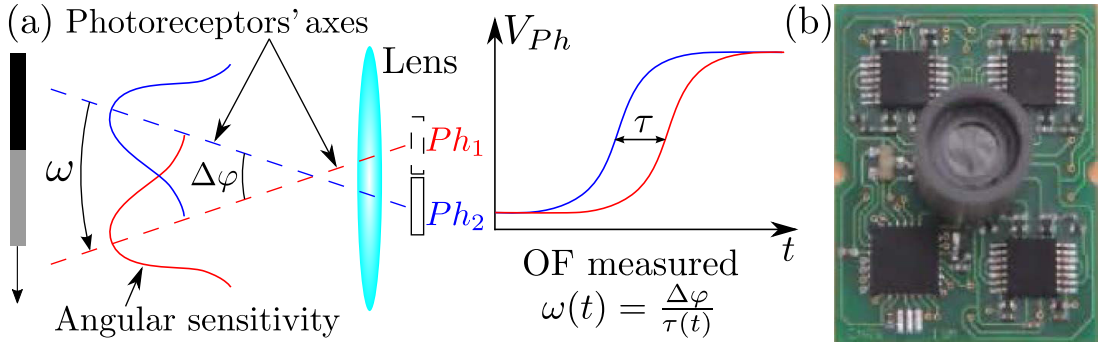

Figure 2.1: (a) Principle of a 2-pixel LMS. The OF produced by a contrast moving in front of the LMS can be computed taking $\omega(t) = \frac{\Delta\varphi}{\tau(t)}$, where $\Delta\varphi$ is the interreceptor angle and τ is the time lag between the two photoreceptors' output signals [122]. (b) Photo of a 6-pixel LMS. (Adapted from [216]).

A defocused lens placed in front of 2 photoreceptors determines the interreceptor angle $\Delta\varphi$ between the 2 photoreceptors' axes and gives them a Gaussian angular sensitivity, on similar lines to what occurs in many insects' eyes. The width of the angular sensitivity determines the photoreceptors' FOV.

A visual contrast moving in front of the LMS will induce a time lag τ between the photoreceptors' output signals. By measuring this time lag, the optical flow can be computed as follows: $\omega(t) = \frac{\Delta\varphi}{\tau(t)}$ [122].

Let us take the case where one N-pixel LMS is installed at each of the four corners of the vehicle, giving up to $4(N-1)$ 1-D OF measurements all round the car. In order to obtain a 180^o FOV of each LMS, N is taken to be such that $N \times \Delta\varphi = 180^o$, where $\Delta\varphi$ is the interreceptor angle, as shown in Fig. 2.1. As it is difficult in practice to obtain a 180^o FOV with just one lens, there could be M N-pixel LMSs, so that $M \times N \times \Delta\varphi = 180^o$.

Figure 2.2 gives a diagram of the visual-motion sensor setup on a vehicle, in the case where 4 neighboring visual signals give 3 1-D OF measurements.

If the environment is relatively smooth, i.e. there are only small variations in the distance between the visual patterns and the LMS in the photoreceptors' FOV, we can assume that the photoreceptors will detect visual points on their axes instead of patterns in their FOV. When an OF measurement ω_i is delivered at the time $\bar{t}$, this will therefore be just as if a visual point has moved from the $i-1$-th to the i-th photoreceptor's axis, and this point is now (at $t = \bar{t}$) on the i-th axis.

In this preliminary study, we assumed that the neighboring visual points detected formed straight lines corresponding to the simplified 2-D profiles of the surrounding parked vehicles.

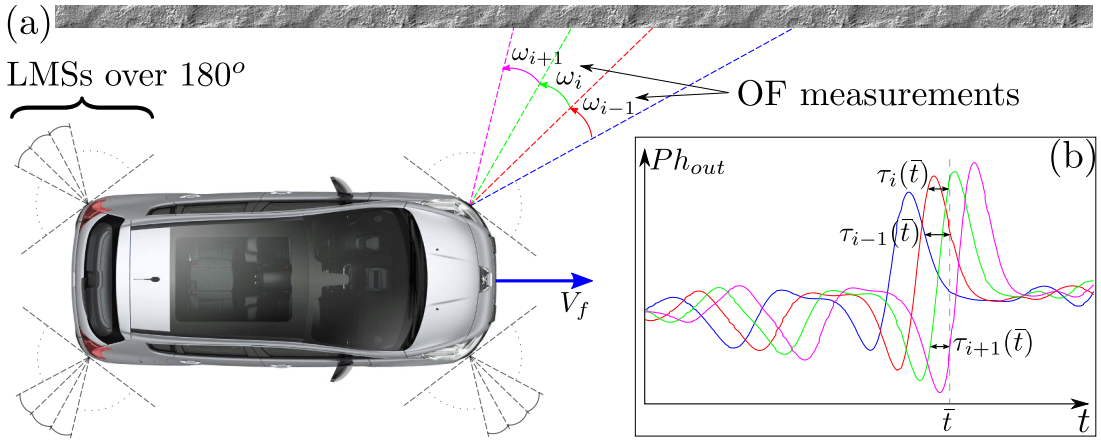

Figure 2.2: (a) Diagram of the visual-motion sensor setup on a vehicle. The N-pixel LMSs are placed at the four corners of the vehicle, where N is such that $N \times \Delta\varphi = 180^o$. (b) An example where 4 neighboring visual signals give 3 local OF measurements.

2.2.4 Point and line motion modeling

As the vehicle's velocity is relatively low during parking maneuvers, we focused here on the 2-D kinematic model for a car-like vehicle moving on a plane.

Figure 2.3 is a top-view diagram of a car-like vehicle with a 1-D OF sensor installed on one corner.

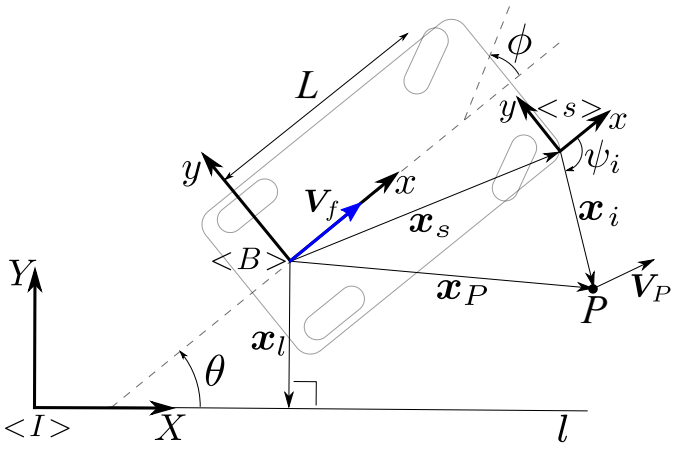

Figure 2.3: Top-view diagram of a car-like vehicle with a 1-D OF sensor installed on one corner.

According to Fig. 2.3 and the Ackermann steering geometry for car-like vehicles, the equations for a point P moving on the same plane as the vehicle can be written as follows:

$$\begin{bmatrix} \dot{x}_i \\ \dot{y}_i \end{bmatrix} = \boldsymbol{f}_P(\boldsymbol{x}_i, \boldsymbol{u}) = \begin{bmatrix} \frac{(y_i+y_s)\tan\phi - L}{L} V_f + V_{Px} \\ -\frac{(x_i+x_s)\tan\phi}{L} V_f + V_{Py} \end{bmatrix} \tag{2.1}$$

$$\omega_i = \frac{x_i\dot{y}_i - y_i\dot{x}_i}{x_i^2 + y_i^2} \tag{2.2}$$

where $\boldsymbol{u} = [V_f\ \phi]^T$ are the vehicle's longitudinal velocity and its steering angle, respectively; L is the

distance between the rear and front wheel axes; ${}^B\boldsymbol{x}_s = [x_s\, y_s]^T$ is the position vector of the 1-D OF sensor with respect to the body frame $< B >$; ${}^B\boldsymbol{x}_P = [x_P\, y_P]^T$, ${}^B\boldsymbol{x}_i = [x_i\, y_i]^T$ are the position vectors of the point P with respect to the body frame $< B >$ and the sensor frame $< s >$, respectively; ${}^B\boldsymbol{V}_P = [V_{Px}\, V_{Py}]^T$ is the velocity vector of P with respect to the inertial frame $< I >$; and ω_i is the angular velocity of P with respect to $< s >$, i.e. the OF measured between the i-th and $i-1$-th pixel of the sensor (see Sec. 2.2.3). Note that the sensor frame $< s >$ is parallel to the body frame $< B >$ (i.e. the rotation matrix between the two frames is the identity matrix), and the upper left index B indicates that the vectors are projected onto the body frame $< B >$. For the sake of simplicity, this index will be dropped in what follows.

As P must lie on the i-th pixel's axis (see Sec. 2.2.3), its coordinates x_i, y_i must satisfy the following equation:

$$\sin\psi_i x_i - \cos\psi_i y_i = 0, \tag{2.3}$$

where ψ_i is the orientation of the i-th pixel's axis, i.e. the angle known to exist between ${}^B\boldsymbol{x}_i$ and the x-axis of $< s >$.

A straight line in the body frame $< B >$ can be described by the following equation:

$$x_l x + y_l y = x_l^2 + y_l^2, \tag{2.4}$$

where ${}^B\boldsymbol{x}_l = [x_l\, y_l]^T$ is the position vector perpendicular to the line l projected onto the body frame $< B >$.

If the line l does not move with respect to the inertial frame $< I >$, we can assume that l is located on the X-axis of $< I >$. In this case, in line with Fig. 2.3, the following equations hold:

$$\begin{bmatrix} \dot{x}_l \\ \dot{y}_l \\ \dot{\theta} \end{bmatrix} = \boldsymbol{f}_l(\boldsymbol{x}_l, \boldsymbol{u}) = \begin{bmatrix} \left(\frac{y_l \tan\phi}{L} - \sin^2\theta\right) V_f \\ -\left(\frac{x_l \tan\phi}{L} + \sin\theta\cos\theta\right) V_f \\ \frac{\tan\phi}{L} V_f \end{bmatrix}, \tag{2.5}$$

where θ is the angle between the X-axis of $< I >$ and the x-axis of $< B >$.

2.2.5 Parking-spot detection and tracking

The algorithm presented here (see Figure 2.4) first computes the positions of the points detected by the LMSs using only 1-D OF measurements and the vehicle's velocity and steering angle (Step 1), and it then looks for straight lines in the clouds of points using the RANSAC method [262] (Step 2). These lines are then classified by Naive Bayes Classifiers (NBCs) (Step 3) to ensure a geometrical coherence with respect to the simplified parking-spot geometry. Lastly, the vectors perpendicular to the parking lines and the intersection points between these lines are estimated by means of Extended Kalman Filters (EKFs) (Steps 4 and 5). A nonlinear tracking control is eventually performed based on the estimation of the vehicle's ego-position with respect to the parking lines and corners, but we will not focus on this process here.

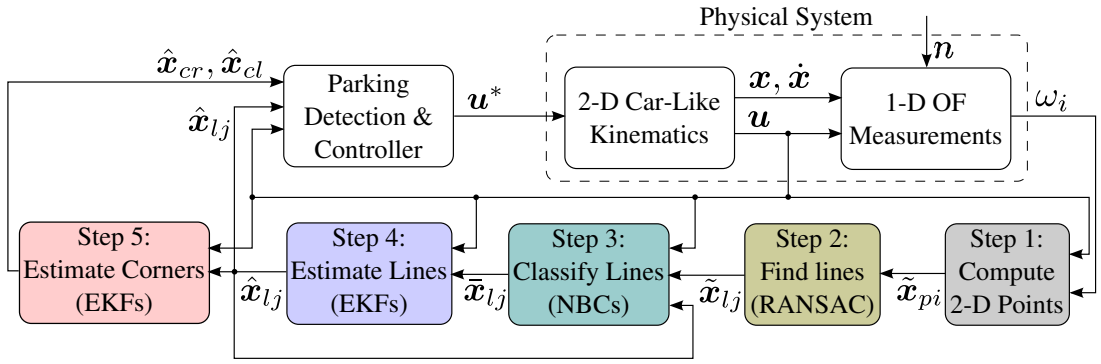

Figure 2.4: Block diagram of the parking-spot detection and tracking algorithm.

Figure 2.5 shows an example of the results obtained in each step in the algorithm at one sampling interval during the simulation. This example involves the case of perpendicular parking, but the present algorithm can be applied to the case of parallel parking as well without making any changes.

In the simulated 2-D parking environment, the parked vehicles were simulated in the form of rectangular shapes (Fig. 2.5(a)) and the OF measurements of the LMSs were computed analytically as the angular velocity of the points of intersection between the lines simulating the parked vehicles and the pixels' axes. A white Gaussian noise ($\boldsymbol{n}$) with the standard deviation $\sigma_{\boldsymbol{n}}$ is added to the positions of the intersections in order to obtain a non-linear noise in the OF measurements, which resembles the real sensors' noise more closely.

Step 1 - Computing 2-D Points

Let us now assume that the point P in Fig. 2.3 does not move with respect to $< I >$ or that its velocity components (V_{Px}, V_{Py}) are negligible with respect to the visual motion induced by the vehicle's motion, so that we can take $\boldsymbol{V}_P \approx \mathbf{0}$.

If we know the vehicle's longitudinal velocity and its steering angle (V_f, ϕ), it is possible by combining equation (2.1), (2.2) and (2.3) to compute the P coordinates as follows[3]:

$$\psi_i \neq \pm\frac{\pi}{2} : \begin{cases} x_i = \frac{V_f(L\tan\psi_i - x_s\tan\phi - y_s\tan\psi_i\tan\phi)}{(\tan^2\psi_i+1)(L\omega_i+V_f\tan\phi)} \\ y_i = x_i\tan\psi_i \end{cases} \tag{2.6}$$

$$\psi_i = \pm\frac{\pi}{2} : \begin{cases} x_i = 0 \\ y_i = \frac{V_f(L - y_s\tan\phi)}{L\omega_i+V_f\tan\phi} \end{cases} \tag{2.7}$$

Figure 2.5(b) gives the cloud of points computed at one sampling interval during the simulation.

[3]Substituting (2.1) into (2.2) is possible because the sampling period Δt is low enough with respect to the dynamics of x_i, y_i. However, a similar result can be obtained in cases where this condition is not met, by taking a discrete approximation of (2.1).

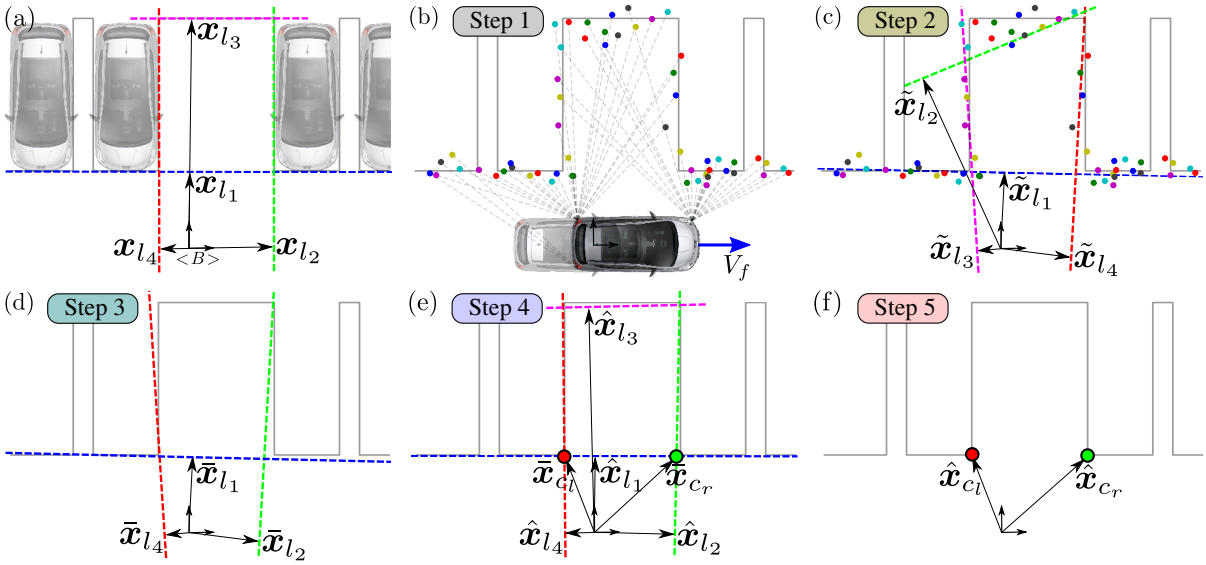

Figure 2.5: Example of results obtained in each step in the parking-spot detection and tracking algorithm at one sampling interval during the simulation. (a) The simplified 2-D parking spot with the position vectors perpendicular to the 4 lines defining the spot. (b) The points detected on the pixels' axes, the positions of which were computed using the 1-D OF measurements delivered by the LMSs placed at the vehicle's corners (Step 1). A white Gaussian noise is added to the positions of these points in order to add a non-linear noise to the OF measurements. (c) The position vectors of the lines obtained by RANSAC in the cloud of points (Step 2). (d) The position vectors of the lines classified by NBCs (Step 3). (e) The position vectors of the lines estimated by EKFs (Step 4). A first approximation of the parking corners' position is given by the points of intersection between each pair of perpendicular lines. (f) The position vectors of the parking corners estimated by EKFs (Step 5).

Step 2 - Finding Lines

To detect a parking spot, first we want to find four straight lines fitting some subsets of the cloud of points obtained in Step 1. Then, to validate this spot, these four lines must satisfy two main conditions: they must be nearly perpendicular two by two, and the distance between the two lines in the two pairs of parallel lines must be greater than the vehicle's width and its length, respectively (see Fig. 2.5(a)). The validation procedure may depend on the specific parking environment (e.g. on whether it is parallel or perpendicular) and require some information from other sensors, and this part will therefore not be discussed here as it goes beyond the scope of this paper.

To look for more than one regression line in a cloud of points, we have to perform some linear clustering in order to obtain subsets containing only one line each.

In this study, we used a modified version of the RANdom SAmple Consensus (RANSAC) method [262], as it provides outliers in a statistically robust way, giving a good compromise between the level of performance and the computational cost with subsets which may be dependent but are not very numerous and are clearly separated linearly (as in the case of the 4 parking lines dealt with here). As previously done

in many vision-based applications such as those designed for motion segmentation [263], the generic RANSAC algorithm is applied recursively to the subset of outliers, i.e., to all the points that do not belong to the consensus set, and whenever a line is detected, it yields the position vector $\tilde{\boldsymbol{x}}_l = [\tilde{x}_l\ \tilde{y}_l]^T$ satisfying equation (2.4). The recursion is stopped when either four lines are detected or no lines are detected in the subset of outliers, i.e., no consensus set is obtained in a maximum number of iterations n. It is worth noting that in order to maintain the high frequency provided by LMSs, we performed only a small number of iterations ($n = 100$). The robustness of the whole algorithm is in fact increased by applying optimal classification (Step 3) and estimation (Step 4,5) methods to the straight lines obtained.

At each sampling interval, the h-th line $\tilde{\boldsymbol{x}}_{l_h}$ given by the algorithm can correspond to any of the four lines defining the parking spot ($\boldsymbol{x}_{l_{1\dots4}}$), or even to none of them.

Figure 2.5(c) shows the four lines detected by RANSAC in the cloud of points presented in Fig. 2.5(b). Note that $\tilde{\boldsymbol{x}}_{l_2}$ (the green line in Fig. 2.5(c)) differs from the corresponding line $\boldsymbol{x}_{l_3}$ (the pink line in Fig. 2.5(a)), while $\tilde{\boldsymbol{x}}_{l_3}, \tilde{\boldsymbol{x}}_{l_4}$ (the red and pink lines in Fig. 2.5(c)) are not in the right order as they correspond to $\boldsymbol{x}_{l_2}, \boldsymbol{x}_{l_4}$ (the green and red lines in Fig. 2.5(a)).

Step 3 - Classifying Lines

Let us take the four straight lines $\boldsymbol{x}_{l_{1\dots4}}$ defining the parking spot (Fig. 2.5(a)) and consider that they belong to four classes $\mathcal{L}_{1\dots4}$.

To determine whether the h-th line $\tilde{\boldsymbol{x}}_{l_h}$ detected in Step 2 belongs to the j-th class $\mathcal{L}_j$, we used Naive Bayes Classifiers (NBCs), giving normal and uncorrelated probability distributions for the two components of $\tilde{\boldsymbol{x}}_{l_h}$. We therefore computed the Bayes probability as follows[4]:

$$\begin{cases} p_{j,h} = p(\mathcal{L}_j | x_{l_h}, y_{l_h}) = \frac{p(\mathcal{L}_j) p(x_{l_h}|\mathcal{L}_j) p(y_{l_h}|\mathcal{L}_j)}{p(x_{l_h}, y_{l_h})} \\ p(x_{l_h}|\mathcal{L}_j) = \frac{1}{\sqrt{2\pi\sigma_{x_j}^2}} e^{-\frac{(x_{l_h} - \mu_{x_j})^2}{2\sigma_{x_j}^2}} \quad \text{(same for } y_{l_h}\text{)} \\ p(x_{l_h}, y_{l_h}) = \sum_{j=1}^{4} p(\mathcal{L}_j) p(x_{l_h}|\mathcal{L}_j) p(y_{l_h}|\mathcal{L}_j) \end{cases} \tag{2.8}$$

where we considered $p(\mathcal{L}_j) = \frac{1}{4}$, $[\mu_{x_j}\ \mu_{y_j}]^T = \hat{\boldsymbol{x}}_{l_j}^-(k)$ and $[\sigma_{x_j}\ \sigma_{y_j}]^T = \boldsymbol{f}_l(\hat{\boldsymbol{x}}_{l_j}(k-1), \boldsymbol{u}(k-1))$, with k denoting the k-th sampling period (i.e., $t = k\Delta t$), and $\hat{\boldsymbol{x}}_{l_j} = [\hat{x}_{l_j}\ \hat{y}_{l_j}]^T$ being the position vector of j-th line estimated by the EKF and $\hat{\boldsymbol{x}}_{l_j}^-$ its "a priori" estimate (see the next subsection for details).

We classified the line $\tilde{\boldsymbol{x}}_{l_h}$ in the class $\mathcal{L}_j$ if $p_{j,h} = \max_{j,h} p_{j,h}$ and $p_{j,h} > 0.95$. To avoid any misunderstandings, we specified the j-th classified line $\bar{\boldsymbol{x}}_{l_j} = [\bar{x}_{l_j}\ \bar{y}_{l_j}]^T$.

Figure 2.5(d) shows the three lines which were classified out of the four lines in Fig. 2.5(c). In particular, $\tilde{\boldsymbol{x}}_{l_2}$ (the green line in Fig. 2.5(c)) was dropped because it was too different from the corresponding line $\boldsymbol{x}_{l_3}$ (the pink line in Fig. 2.5(a)), while $\tilde{\boldsymbol{x}}_{l_3}, \tilde{\boldsymbol{x}}_{l_4}$ (the red and pink lines in Fig. 2.5(c)) were classified as $\bar{\boldsymbol{x}}_{l_4}, \bar{\boldsymbol{x}}_{l_2}$ (the green and red lines in Fig. 2.5(a)).

[4] The hypothesis that the two components of $\tilde{\boldsymbol{x}}_{l_h}$ had normal and uncorrelated probability distributions was adopted on the basis of what was observed statistically during several simulations with trajectories of various kinds.

Step 4 - Estimating Lines

In order to obtain a robust continuous estimation of the parking line vectors, an Extended Kalman Filter (EKF) based on a discrete approximation for the model presented in (2.5) was implemented for each of the four lines $\boldsymbol{x}_{l_{1\ldots4}}$ by taking the lines classified in Step 3 ($\bar{\boldsymbol{x}}_{l_{1\ldots4}}$) to be actual measurements.

For each line, we took the first-order discrete approximation for the model presented in (2.5) as follows:

$$\begin{cases} \boldsymbol{x}_{l_k} = \hat{\boldsymbol{f}}_l(\boldsymbol{x}_{l_{k-1}}, \boldsymbol{u}_{k-1}, \boldsymbol{w}_{l_{k-1}}) \\ \qquad = (\boldsymbol{f}_l(\boldsymbol{x}_{l_{k-1}}, \boldsymbol{u}_{k-1}) + \boldsymbol{w}_{l_{k-1}})\Delta t + \boldsymbol{x}_{l_{k-1}} \quad , \\ \boldsymbol{z}_{l_k} = \boldsymbol{h}_l(\boldsymbol{x}_{l_k}, \boldsymbol{\nu}_{l_k}) = \boldsymbol{x}_{l_k} \end{cases} \tag{2.9}$$

where the index k denotes the k-th sampling period (i.e., $t = k\Delta t$); $\boldsymbol{w}_l, \boldsymbol{\nu}_l$ denotes the model and the measurement noise, respectively, and they are assumed to be independent white noises and to have normal probability distributions, i.e. $p(w) \sim N(0, Q)$ and $p(\nu) \sim N(0, R)$, where Q, R are the covariance matrices which, in the present case, are diagonal[5].

At each sampling period, the steps involved in each EKF can be summarized as follows:

1. $\hat{\boldsymbol{x}}_{l_k}^- = \hat{\boldsymbol{f}}_l(\hat{\boldsymbol{x}}_{l_{k-1}}, \boldsymbol{u}_{k-1}, \boldsymbol{0})$

2. $P_k^- = F_k P_{k-1} F_k^T + W_k Q W_k^T$

3. $K_k = P_k^- H_k^T (H_k P_k^- H_k^T + V_k R V_k^T)^{-1}$

4. $\hat{\boldsymbol{x}}_{l_k} = \hat{\boldsymbol{x}}_{l_k}^- + K_k(\boldsymbol{z}_{l_k} - \boldsymbol{h}_l(\hat{\boldsymbol{x}}_{l_k}^-, \boldsymbol{0}))$

5. $P_k = (I - K_k H_k) P_k^{-1}$

where $F = \frac{\partial \hat{\boldsymbol{f}}_l}{\partial \boldsymbol{x}_l}, H = \frac{\partial \boldsymbol{h}_l}{\partial \boldsymbol{x}_l}, W = \frac{\partial \hat{\boldsymbol{f}}_l}{\partial \boldsymbol{w}_l}, V = \frac{\partial \boldsymbol{h}_l}{\partial \boldsymbol{\nu}_l}$ are the Jacobian matrices; P, P^- are the error covariance matrix and its "a priori" estimate, respectively; and I is the identity matrix.

In the present study, the j-th EKF was enabled when the j-th line was detected and validated for the first time on the basis of the time correlation in the values of $\tilde{\boldsymbol{x}}_{l_j}$ (hypothesis test), but in practice, this could be done using other sensors such as ultrasonic sensors, cameras or radars installed on the vehicle.

The initial estimate of $\hat{\boldsymbol{x}}_{l_j}$ was set at the first value of $\tilde{\boldsymbol{x}}_{l_j}$ validated, while the initial estimate of P_k was set at the identity matrix.

When there was no measurement of the j-th EKF, i.e. no $\bar{\boldsymbol{x}}_{l_j}$ in Step 3, the Kalman gain K_k was set at zero, so that it was still possible to have an estimation of $\boldsymbol{x}_{l_j}$ based on the "a priori" prediction $\hat{\boldsymbol{x}}_{l_j}^-$.

Figure 2.5(e) shows the four lines estimated by the four EKFs. Note that although there was no measurement corresponding to $\bar{\boldsymbol{x}}_{l_3}$ in Fig. 2.5(d), we still had an estimation of $\hat{\boldsymbol{x}}_{l_3}$.

[5] The hypothesis that the components of the model and the measurement noise had normal and uncorrelated distributions was adopted on the basis of what was observed statistically during several simulations with trajectories of various kinds.

Step 5 - Estimating Corners

The position vectors of the four parking corners can be estimated in terms of the intersection point between each pair of perpendicular lines estimated in Step 4. This makes it possible to compute the j-th corner vector $\bar{\boldsymbol{x}}_{c_j} = [\bar{x}_{c_j}\, \bar{y}_{c_j}]^T$ by substituting $\hat{\boldsymbol{x}}_{l_j}$ and $\hat{\boldsymbol{x}}_{l_{j+1}}$ into (2.4) and combining the two resulting equations. Here we dealt only with the two outer corners (the green and red circles in Fig. 2.5(e)). We took $\bar{\boldsymbol{x}}_{c_l}, \bar{\boldsymbol{x}}_{c_r}$ to specify the position vectors of the left and right corners, respectively.

This estimate of the corners' positions was not accurate enough to be able to perform a feedback tracking control on it. In particular, when the car is crossing the lines, i. e. $\boldsymbol{x}_l \approx \mathbf{0}$, a very small error in $\hat{\boldsymbol{x}}_l$ can lead to a very large error in $\bar{\boldsymbol{x}}_c$.

In order to obtain a robust continuous estimation of the corner vectors, an EKF based on a discrete approximation of the model presented in (2.1) was implemented for each of the two outer corners $(\boldsymbol{x}_{c_l}, \boldsymbol{x}_{c_r})$, by taking the line intersection points defined above $(\bar{\boldsymbol{x}}_{c_l}, \bar{\boldsymbol{x}}_{c_r})$ to be actual measurements.

The same procedure as that applied with the EKF on $\boldsymbol{x}_l$ in the previous subsection can be used with the EKF on $\boldsymbol{x}_c$ (taking the same model and the same measurement noise). We only had to replace the index l by the index c and take the function $\boldsymbol{f}_P(\boldsymbol{x}_c, \boldsymbol{u})$ defined in (2.1) instead of $\boldsymbol{f}_l(\boldsymbol{x}_l, \boldsymbol{u})$.

The EKFs for $\boldsymbol{x}_{c_l}$ and $\boldsymbol{x}_{c_r}$ were enabled when the EKFs for $\boldsymbol{x}_{l_1}, \boldsymbol{x}_{l_2}$ and $\boldsymbol{x}_{l_1}, \boldsymbol{x}_{l_4}$ were enabled, respectively (the blue and green lines and the blue and red lines, respectively, in Fig. 2.5(e)).

The initial estimates for $\hat{\boldsymbol{x}}_{c_l}$ and $\hat{\boldsymbol{x}}_{c_r}$ were set at the first value of the intersection point between $\hat{\boldsymbol{x}}_{l_1}, \hat{\boldsymbol{x}}_{l_2}$ and $\hat{\boldsymbol{x}}_{l_1}, \hat{\boldsymbol{x}}_{l_4}$, respectively, while the initial estimate of P_k was set at the identity matrix.

Note that although we defined the parking spot by the above four lines $\boldsymbol{x}_{l_{1\ldots4}}$, the parking spot can be detected even if x_{l_3} is not found (i.e. in the case of an "open" or "homogeneous" background), and depending on the automatic pilot's requirements, a tracking control of the ego-vehicle's trajectory can eventually be performed only on the basis of the two outer corners' positions $(\boldsymbol{x}_{c_l}, \boldsymbol{x}_{c_r})$ and the two lateral lines' orientations $(\boldsymbol{x}_{l_2}, \boldsymbol{x}_{l_4})$. For safety reasons, information about the parking-spot depth (i.e. the distance between the vehicle and $\boldsymbol{x}_{l_3}$) should be provided by means of other sensors (e.g. ultrasonic sensors or cameras).

2.2.6 Closed-loop simulation results

In the simulation presented here, we used 40-pixel LMSs with $\Delta\varphi = 4.5^o$ $(N \times \Delta\varphi = 40 \times 4.5^o = 180^o)$ and a white Gaussian noise with standard deviation $\sigma_{\boldsymbol{n}} = 10^{-2}$. The algorithm was run at $100\,Hz$, giving a sampling period of $\Delta t = 10\,ms$. We decided to simulate the case of perpendicular parking because it involves more complex maneuvers than parallel parking, which gives noisier OF measurements than the parallel case. In addition, to obtain results that are closer to the real case, we took into account only the OF measurements in $[-350°, -1°]$ and $[1°, 350°]$ (see [216, 242]).

For the simulated results presented here, the stages of the parking controller can be summarized as follows:

Parking stage 1 - V_f constant, ϕ in closed loop

Drive at constant velocity ($V_f = 1\,\frac{m}{s}$) parallel to the first parking line x_{l_1} (blue line in Fig. 2.5) until the two outer corners $\boldsymbol{x}_{c_l}, \boldsymbol{x}_{c_r}$ are found: if the distance between $\boldsymbol{x}_{c_l}$ and $\boldsymbol{x}_{c_r}$ is greater than a minimum value of the parking-spot width then go to the next step, otherwise continue to look for a parking spot. To drive parallel to x_{l_1}, the steering angle was computed as $\phi = \arctan(k_\phi L \frac{\theta^* - \hat{\theta}_{l_1}}{V_f})$, where $\hat{\theta}_{l_1}$ is the vehicle's orientation with respect to x_{l_1} estimated by the first EKF (see equation (2.9)), $\theta^* = \frac{\pi}{2}$ and $L = 2\,m$.

Parking stage 2 - V_f and ϕ constant

Drive away from the parking spot at constant velocity and steering angle ($V_f = 1\,\frac{m}{s}$, $\phi = -\frac{\pi}{9}$) for a fixed time in order to bring the vehicle to a suitable pose for facilitating the backwards maneuver. Then, compute the reference trajectory $\boldsymbol{X}_c^*$ of the vehicle's ego-position with respect to the inertial frame $< I >$ centered to one of the two corners. The trajectory from the actual pose to the desired pose, i.e. being in the middle between $\boldsymbol{x}_{c_l}$ and $\boldsymbol{x}_{c_r}$ and parallel to $\boldsymbol{x}_{l_2}$ (or $\boldsymbol{x}_{l_4}$), was computed as a third-order polynomial function using the boundary conditions on the initial and final position and velocity.

Parking stage 3 - V_f and ϕ in closed loop

Drive backwards following the reference trajectory by applying a feedback linearizing control based on the Ackermann model as follows [264]: $V_f = -\sqrt{\dot{X}_c^{\,2} + \dot{Y}_c^{\,2}}$ and $\phi = \arctan(L \frac{\dot{X}_c \ddot{Y}_c - \dot{Y}_c \ddot{X}_c}{V_f^3})$, where $\dot{\boldsymbol{X}}_c = [\dot{X}_c\, \dot{Y}_c]^T = \dot{\boldsymbol{X}}_c^* + k_1(\boldsymbol{X}_c^* - \hat{\boldsymbol{X}}_{c_l})$ and $\ddot{\boldsymbol{X}}_c = [\ddot{X}_c\, \ddot{Y}_c]^T = \ddot{\boldsymbol{X}}_c^* + k_2(\dot{\boldsymbol{X}}_c^* - \dot{\hat{\boldsymbol{X}}}_{c_l})$, $\hat{\boldsymbol{X}}_{c_l}$ being the estimation on the vehicle's ego-position with respect to the left corner, i.e. $\hat{\boldsymbol{X}}_{c_l} = -{}^I R_B \hat{\boldsymbol{x}}_{c_l}$, and ${}^I R_B$ being the rotation matrix from $< B >$to $< I >$.[6]

Parking stage 4 - V_f constant , ϕ in closed loop

Drive backwards at low constant velocity ($V_f = -0.5\,\frac{m}{s}$) parallel to $\boldsymbol{x}_{l_2}$ (or $\boldsymbol{x}_{l_4}$) by computing the steering angle as $\phi = \arctan(k_\phi L \frac{\theta^* - \hat{\theta}_{l_2}}{V_f})$, with $\theta^* = -\frac{\pi}{2}$, until the parking maneuver is done.[7]

Figure 2.6 shows the errors in the line and corner position vectors which occurred in Steps 2 to 5, as described in Sec. 2.2.5, during a given trajectory of the simulated vehicle.

In Step 2, the errors between the real lines and the lines obtained with RANSAC ($\boldsymbol{x}_l - \tilde{\boldsymbol{x}}_l$) were sometimes relatively large (Fig. 2.6(a)). As explained above, this was due to the fact that the lines given by RANSAC were sometimes not in the right order (for instance, the second line obtained occasionally corresponded to the third one, as in Fig. 2.5(c)), or they were sometimes spurious lines.

In Step 3, the errors between the real lines and the lines classified by NBCs ($\boldsymbol{x}_l - \bar{\boldsymbol{x}}_l$) were much smaller, but they could still be relatively large, depending on the noise (Fig. 2.6(b)).

[6] We used the estimation $\hat{\boldsymbol{x}}_{c_l}$ (left corner) since the estimation $\hat{\boldsymbol{x}}_{c_r}$ (right corner) was not enough reliable during the first part of the maneuver because we have no measurements for x_{l_2} due to the occlusion.

[7] The gains $k_\phi, k_1, k_2 > 0$ were dynamically adapted in order to avoid any saturation of the control inputs.

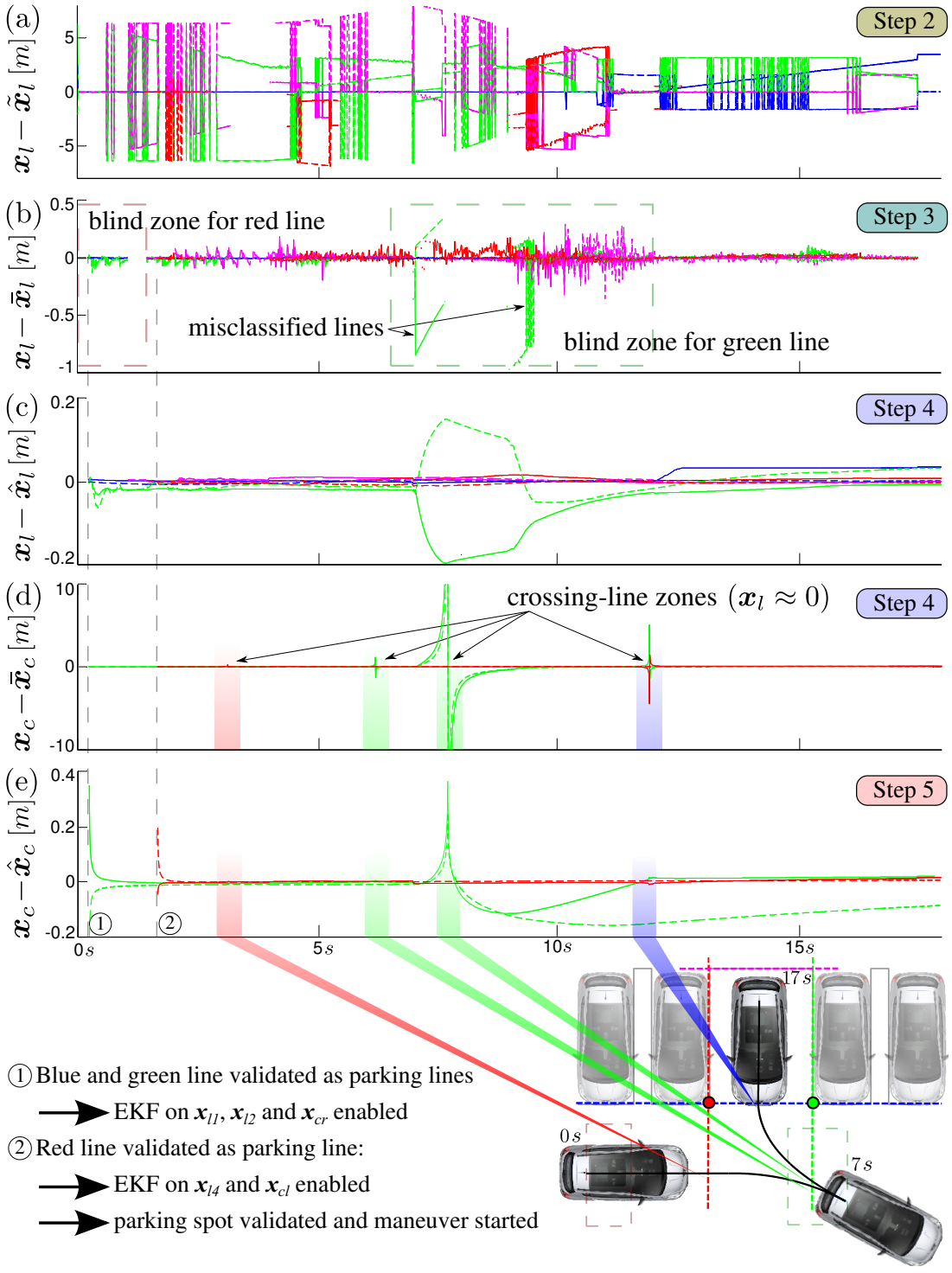

Figure 2.6: Errors in the line and corner position vectors occurring during a given trajectory of the simulated vehicle at: (a) Step 2 ($\boldsymbol{x}_l - \tilde{\boldsymbol{x}}_l$); (b) Step 3 ($\boldsymbol{x}_l - \bar{\boldsymbol{x}}_l$); (c) Step 4 ($\boldsymbol{x}_l - \hat{\boldsymbol{x}}_l$); (d) Step 4 ($\boldsymbol{x}_c - \bar{\boldsymbol{x}}_c$); (e) Step 5 ($\boldsymbol{x}_c - \hat{\boldsymbol{x}}_c$). The colors in this figure correspond to those of the lines and corners presented in Fig. 2.5. The continuous and dotted curves correspond to the x and y coordinates of the vectors, respectively.

In Step 4, the errors between the real lines and the lines estimated by the EKFs $(\boldsymbol{x}_l - \hat{\boldsymbol{x}}_l)$ were about one order of magnitude smaller, and an estimation was still delivered even when no corresponding line was found, i.e. no measurement $\bar{\boldsymbol{x}}_l$ was obtained (Fig. 2.6(c)). However, when the car was crossing the lines, i. e. $\boldsymbol{x}_l \approx \mathbf{0}$, a very small error in $\hat{\boldsymbol{x}}_l$ could lead to a very large error in $\bar{\boldsymbol{x}}_c$ (corresponding to the shaded regions in Fig. 2.6(d)).

In Step 5, the errors between the real corners and those estimated by the EKFs $(\boldsymbol{x}_c - \hat{\boldsymbol{x}}_c)$ were much smaller, especially when crossing lines (Fig. 2.6(e)).

The preliminary controller and the results presented here could be further improved, for instance by looking for the right corner's position $\boldsymbol{x}_{c_r}$ in the cloud of points while the vehicle is in the blind zone for the x_{l_2} (see Fig. 2.6) in order to improve the reliability on the estimation $\hat{\boldsymbol{x}}_{c_r}$ and perform a closed-loop maneuver based on the estimation of both corners' positions.

2.2.7 Preliminary experimental tests

Preliminary tests were carried out on a real vehicle (*Peugeot 3008*) in order to check the validity of the first step in the algorithm presented in Sec. 2.2.5. In particular, two 6-pixel LMSs ($\Delta\varphi \approx 1.5^o$ [216] and 3.8^o [242]) and a webcam (*Logitech B905*) were tied to a supporting slab which was attached to the front fender, as shown in Fig. 2.7(a). The slab could be moved and rotated along the z-axis in order to adjust the sensors' height and their orientation.

Figure 2.7(b) shows an example of the OF measured by one of the two 6-pixel LMSs presented in Fig. 2.7(a) ($\Delta\varphi \approx 3.8^o$ [242]) when the vehicle was moving alongside two parallel parked vehicles. Figure 2.7(c) shows the vehicle's trajectory, along with the positions of the 2-D points computed by applying equation (2.6) to the OF measurements in Fig. 2.7(b).

To compute the OF measurements ω_i, the time lag τ_i between the i-th and $i-1$-th pixel output signal (see Fig. 2.1, 2.2) was estimated using a cross-correlation method inspired by the Reichardt-Hassenstein model [76]. First we delayed one of the two signals by a time $\tau_k = k\Delta t$ in a fixed time window, and we then computed the cross-correlation between the delayed and non-delayed signals. We then set τ_i at a value equal to the time lag τ_k giving the maximum cross-correlation, as long as this maximum was greater than 0.99. As the precision in τ_i is constant and ω_i is inversely proportional to τ_i, the precision in ω_i is proportional to ω_i itself, as we can notice by looking at Fig. 2.7(b).

The data acquisition and the data processing were performed on a laptop PC (*Intel Core i7-2620M CPU 2.70 GHz*) with RTMaps software. In particular, the pixel output signals delivered by the LMSs were sampled at a rate of $1\,kHz$ by the embedded micro-controller (*Microchip dsPIC33FJ128GP802*) and transmitted to the PC via serial communication, whereas the vehicle's longitudinal velocity and its steering angle were acquired at $20\,Hz$ by the PC via the vehicle's CAN bus. The OF and the positions of the 2-D points were computed at a frequency of $100\,Hz$.

Figure 2.8 shows two examples of the vehicle's trajectory along with the positions of the 2-D points computed using the OF measured by one of the two 6-pixel LMSs presented in Fig. 2.7(a) ($\Delta\varphi \approx 3.8^o$

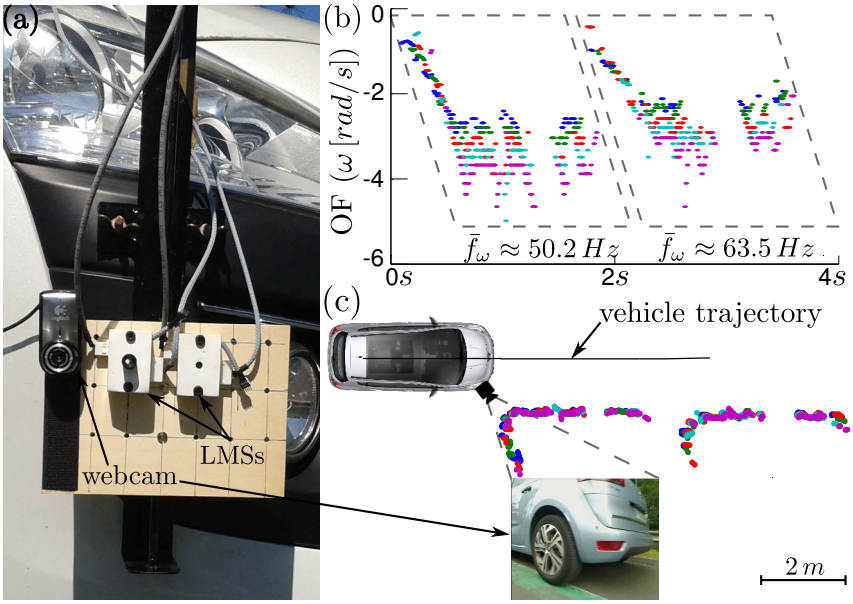

Figure 2.7: (a) A picture of the sensor's supporting slab attached the front fender with the two 6-pixel LMSs and the webcam. (b) Example of the OF measured by one of the two 6-pixel LMSs presented in Fig. 2.7(a) ($\Delta\varphi \approx 3.8^o$ [242]) when the vehicle was moving alongside two parallel parked vehicles. The dotted lines indicate the OFs produced by each of the two parked vehicles, and $\bar{f}_\omega$ indicates the average refresh frequency of the OF measurements in each dotted-line region. (c) The simulated vehicle's trajectory, along with the positions of the 2-D points computed by applying equation (2.6) to the OF measurements presented in Fig. 2.7(b). The picture in the lower part of the figure is the webcam image taken when the vehicle was in the position shown here (webcam's FOV $\sim 45^o$).

[242]) when the vehicle was moving along a parallel (Fig. 2.8(a)) and a perpendicular parking lane (Fig. 2.8(b)) at a longitudinal velocity ranging from 3 to 9 $\frac{m}{s}$.

The refresh frequency of the OF measurements, that is to say the ratio between the number of measurements performed in a given time interval and the time interval itself, can vary significantly, depending on the empty spaces and the contrasts provided by the car bodies (see the pictures in the lower part of Fig. 2.8(a) and (b)), giving both dense and sparse clouds of points. However, the average refresh frequency $\bar{f}_\omega$ for the OF produced by the parked vehicles ranged from 30 to 65 Hz, which is about 2-3 times greater than those of standard ultrasonic sensors and cameras.

A few 2-D points were generated in the center of Fig. 2.8(a) as the vehicle was moving along an empty parking place lined with small traffic poles (see the third picture in Fig. 2.8(a)). No measurements were made here on the background of the spots because it was too far from the LMS. In fact, the analog band-pass filter implemented on the LMS (see [242]) cuts off the low-frequency variations in the photoreceptors' signals produced by the slow visual motion of the background. This issue can be fixed by adjusting the analog filtering part of the LMS, but this has not been shown here.

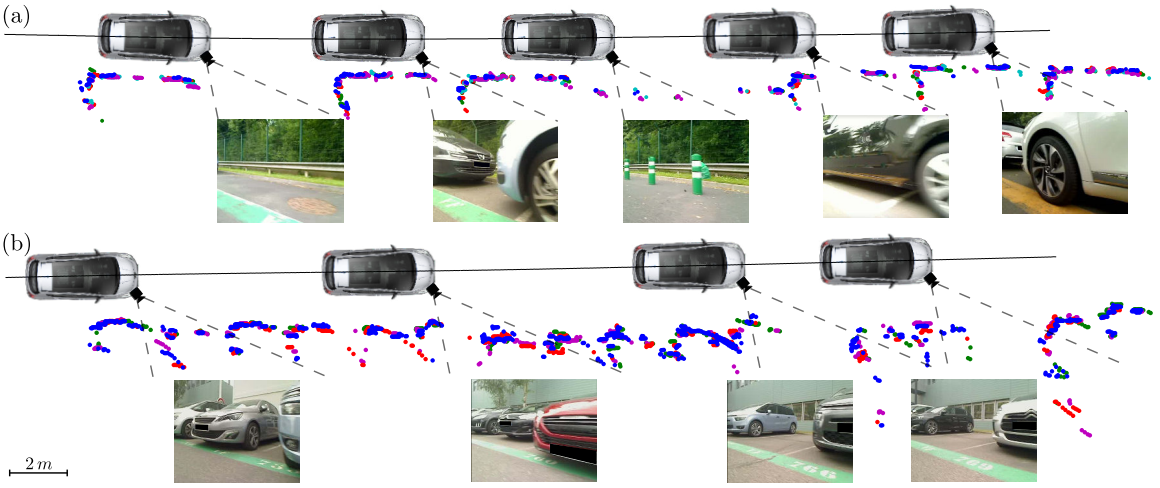

Figure 2.8: Example of vehicle's trajectory along with the positions of the 2-D points computed using the OF measured by one of the two 6-pixel LMSs presented in Fig. 2.7(a) ($\Delta\varphi \approx 3.8^o$ [242]) when the vehicle was moving along a (a) parallel and (b) a perpendicular parking lane at a longitudinal velocity ranging from 3 to 9 $\frac{m}{s}$. The pictures in the lower part of each subfigure are the webcam images taken when the vehicle was in the positions presented in the top part, and the dotted lines give a rough indication of the webcam's FOV ($\sim 45^o$).

It is worth noting that in the case of both parallel and perpendicular parking, it is possible to recognize the shapes of the parked vehicles in the clouds of points, even though few measurements are sometimes made due to the reflections on the car bodies (see, for instance, the fourth picture in Fig. 2.8(a)).

In the case of perpendicular parking, the measurements were noisier than in the parallel case, mainly because of (i) the occlusions of the cars' sides and (ii) the misalignment of the cars (see for instance the first and second pictures in Fig. 2.8(b)).

The qualitative results presented here validate the first step in the algorithm we have developed. Although it was not possible, in the case of this example, to apply the other steps in the algorithm because there were too few points in each sampling interval, it seems likely that by using LMSs with a larger number of pixels (or several 6-pixel LMSs) delivering 1-D OF measurements all round the vehicle, we will probably have enough points to be able to apply all the steps.

2.2.8 Conclusions

In this paper, we have presented a low computational-cost method for detecting and tracking a parking spot in real time based on a visual motion sensor setup performing 1-D OF measurements around the vehicle.

The advantages of this method can be summed up as follows:

- the average refresh frequency is about 2-3 times higher than with standard ultrasonic sensors and cameras;

- it produces redundant information about a parking spot, in cases where ultrasonic sensors and cameras are liable to be inaccurate (e.g., when the parked vehicles are too far apart or too close, under wide-range lighting conditions and when indistinct visual patterns are visible on the car bodies);
- it can detect a candidate parking spot before passing it, thanks to the long-distance, wide-angle range of view;
- it can track the parking spot in real time, thanks to the wide-angle range of view, as well as the high-frequency measurements and low computational-cost processing;
- it makes it possible to eventually apply a nonlinear tracking control strategy using the estimation of the vehicle's ego-position with respect to the parked vehicles.

For these reasons, LMSs can provide a good additional information to automatic parking systems as well as to other automotive applications. In particular, by using LMSs with several pixels (or several few-pixel LMSs) giving a 180^o FOV at each of the car's four corners, it should be possible to apply the present method in real time to performing safe closed-loop parking maneuvers.

Although we focused here only on the information provided by the LMSs, data provided by other sensors (such as ultrasonic sensors) could be integrated into the algorithm to achieve more robust performances.

Studies to improve the closed-loop parking maneuvers and experimental tests with several LMSs on a real car-like robot are now under way.

2.3 Conclusion

As commented at the end of the paper, we showed that LMSs might provide good additional information to automatic parking maneuvers, as well as to other maneuvers in which a wide FOV and high-frequency measurements are needed. In addition, these sensors can deliver redundant information in cases where ultrasonic sensors and cameras are liable to be inaccurate, for instance, when the parked vehicles are too far apart or too close, under wide-range lighting conditions and when poor visual patterns are visible on the car bodies.

We also wanted to propose here a relatively simple method to estimate in real time the vehicle's ego-position using 2-D point clouds giving quasi-rectangular shapes of the parked vehicles instead of using odometry, which may lead to unreliable or inaccurate position estimates. The method we proposed would allow to detect a candidate parking spot before passing it, thanks to the long-distance, wide-angle range of view, and therefore drive the vehicle in a good starting pose and eventually apply an "optimal" trajectory control strategy based on a real-time estimation of the vehicle's ego-position with respect to the parked vehicles. Although we showed that in a simplified simulated parking situation 1-D optic flow measurements are sufficient for performing closed-loop parking maneuvers, this method have some important limitations, e.g. the hypothesis that the parked vehicles nearly do not move, and therefore in a real parking situation data provided by other sensors (e.g. ultrasonic sensors) should be integrated to achieve higher robustness and safety.

The preliminary experimental tests performed in real parking conditions showed that the LMSs developed so far have some drawbacks related to high-dynamic-range lighting conditions, and therefore are not ready yet to integrate ADAS systems. In fact, the response of the pixels used saturated when the light changes were relatively large, which may occur, for instance, when passing from shaded to lightened areas or when direct sunlight is reflected on the cars body, and they had different contrast sensitivities at different light levels (e.g. from night to day time).

In addition, the new cross-correlation method used for computing the optic flow had two limitations: it was not yet adapted for implementation onboard light embedded electronics, and the resolution of the measurements was not constant but depended on the amplitude of the measurements themselves. This last limit is a well-known issue for time-of-travel schemes: since the optic flow is computed as $\omega = \frac{\Delta\varphi}{\tau}$, if the resolution of the time lag τ is linear, i.e. $\tau = n\Delta t$ with Δt being the sampling time and n an integer, then the resolution of the optic flow ω follows an hyperbolic law giving low $\Delta\omega$ at high τ (i.e. n) and viceversa (see Fig. 2.7(b)). Such a hyperbolic resolution may result in a higher noise level in the 2-D point clouds computed from the optic flow measurements.

The next steps were therefore to improve these sensors in order to deliver robust measurements in different lighting conditions, and then optimize the optic flow processing so that it can be implemented on embedded electronics and give a constant OF resolution. To this purpose, first the work carried out on a new auto-adaptive silicon retina will be presented in the next chapter, and then the new optic flow sensor based on these improvements will be presented in chapter 4.

Chapter 3

M^2APix: a Novel Bio-Inspired Auto-Adaptive Pixel

3.1 Introduction

In the previous chapter, we showed that LMSs could be promising for ADAS systems, however some improvements have to be applied to these sensors in order to provide a higher level of robustness. In particular, the LMSs used for the preliminary experiments carried out on a vehicle, which were developed in previous thesis [216, 242], sometimes delivered inaccurate responses depending on the lighting conditions.

As discussed in section 1.8.1, this is a known, common problem for visual sensors that have to operate in high-dynamic-range lighting conditions. Thus, it was not surprising for us to find out that LMSs have to be improved in this sens in order to correctly work outdoors and particularly in road environments. Some solutions using auto-adaptive pixels, namely the Delbruck pixel [175] (see section 1.8.1), have already been presented in [177,242], however their responses were not satisfactory for optic flow measurements when the light changes were higher than 1 decade (see Fig. 7(b),(d),(j),(l) in [177]).

We have therefore worked on a new auto-adaptive silicon retina, which comprises both the pixels of Delbruck type [175] and novel pixels, called M^2APix, that reproduce the Michaelis-Menten law inspired by the findings on the animal and human retina (see section 1.6.1). The challenge of this novel pixel was to improve some limits of existing auto-adaptive pixels (see section 1.8.1): (i) keeping high and constant contrast sensitivity independently to the light level, (ii) not deviating when large changes in light occur due to saturation of the VLSI transistors, and (iii) adapting in a luminosity range as wide as possible. Since it was the first implementation of the M^2APix pixel on a VLSI chip, the Delbruck pixels were included to (i) provide a meaningful comparison of the M^2APix responses with those of a well-known auto-adaptive pixel, as the technology used for both pixels was the same, and (ii) have a useful chip even in the unlucky case in which M^2APix did not deliver a better response.

For this purpose, we also developed a new standard method for accurately characterizing pixels' responses to luminous changes of up to ± 3 decades in a 7-decade mean luminosity range, by implementing a single light source, which has been called the Lighting Box. In fact, in the studies presented in section 1.8.1, the light adaptation and contrast sensitivity of silicon retinas were often tested by applying a series of lighting steps (AC light) in addition to various background lights (DC light). As a result, the pixels' responses have often been described in terms of the stimulus intensity, and direct comparisons could therefore be made with the biological findings. However, the method used to characterize pixels' responses was sometimes not clear or had not even been described at all, which makes comparisons with other results very difficult.

In the paper we present here, we described in details the M^{2}APix pixel, as well as the new method proposed for accurately characterizing pixels' responses, and presented a complete characterization this novel pixel, showing very promising results for use in optic flow-based applications. The VLSI chip was designed and developed by our laboratory in collaboration with the Center of Particle Physics of Marseille (CPPM) prior to this thesis. My personal contribution in this work was the design and development of the hardware and software setup for testing the chip and processing the data, as well as the bibliographic study and the main writing of the paper.

Some additional information on the VLSI chip as well as the characterization results obtained for the Delbruck pixels present in this chip are given in Appendix B.

3.2 Article

"A bio-inspired analog silicon retina with Michaelis-Menten auto-adaptive pixels sensitive to small and large changes in light"

Published in 2015 in *Optics Express*, 23(5), pp. 5614–5635

Authors: Stefano Mafrica,[1, 2] Stéphanie Godiot,[3] Mohsine Menouni,[3] Marc Boyron,[1] Fabien Expert,[1] Raphaël Juston,[1] Nicolas Marchand,[4] Franck Ruffier,[1] Stéphane Viollet[1]

3.2.1 Abstract

In this paper, we present: (i) a novel analog silicon retina featuring auto-adaptive pixels that obey the Michaelis-Menten law, i.e. $V = V_m \frac{I^n}{I^n+\sigma^n}$; (ii) a method of characterizing silicon retinas, which makes it possible to accurately assess the pixels' response to transient luminous changes in a ± 3-decade range, as well as changes in the initial steady-state intensity in a 7-decade range. The novel pixel, called M^{2}APix, which stands for Michaelis-Menten Auto-Adaptive Pixel, can auto-adapt in a 7-decade range and responds appropriately to step changes up to ± 3 decades in size without causing any saturation of the Very Large Scale Integration (VLSI) transistors. Thanks to the intrinsic properties of the Michaelis-Menten equation, the pixel output always remains within a constant limited voltage range. The range of the Analog to Digital Converter (ADC) was therefore adjusted so as to obtain a Least Significant Bit (LSB) voltage of $2.35\,mV$ and an effective resolution of about 9 bits. The results presented here show that the M^{2}APix produced a quasi-linear contrast response once it had adapted to the average luminosity. Differently to what occurs in its biological counterparts, neither the sensitivity to changes in light nor the contrast response of the M^{2}APix depend on the mean luminosity (i.e. the ambient lighting conditions). Lastly, a full comparison between the M^{2}APix and the Delbrück auto-adaptive pixel is provided.

3.2.2 Introduction

During the last few decades, research in the field of robotics has advanced considerably, but there still exist very few sighted robots which are able to behave appropriately, regardless of changes in the illuminance (see Fig. 5 in [242], for instance), such as those which occur outdoors. One of the reasons for this lack is that it is difficult to design pixels that combine high sensitivity with a wide luminosity range.

A large variety of Wide-Dynamic-Range (WDR) CMOS image sensors has been proposed throughout the years [156], trying to widen the operating luminosity range as much as the visible spectrum while

[1] Aix-Marseille Université, CNRS, ISM UMR 7287, 13288 Marseille, France
[2] PSA Peugeot Citroën, Route de Gisy, 78140 Vélizy-Villacoublay, France
[3] Aix-Marseille Université, CNRS/IN2P3, CPPM UMR 7346, 13288 Marseille, France
[4] GIPSA-lab Laboratory, Control Systems Department, SySCo team, CNRS, University of Grenoble, ENSE3 BP 46, 38402 St Martin d'Hères Cedex, France

keeping sensitivity to small changes for every average luminosity in the operating range. Although the WDR image sensors capture images in a luminosity range of up to 7 decades, they provide different contrast sensitivities at different average luminosities. Vision applications such as event-based and biomedical applications often require high and constant sensitivity in a large luminosity range, in order to detect small temporal and/or spatial changes in the intensity in several lighting conditions [169, 170, 265, 266]. One possible solution to this problem can be found by looking at the auto-adaptive response of human and animal photoreceptors.

In their physiological studies on fish, Naka and Rushton established for the first time that a vertebrate retina obeys a process of adaptation whereby each photoreceptor's response is normalized by a representative value of the average local luminosity [36], in line with the Michaelis-Menten equation [44]. In the light of these and subsequent findings [36, 39, 43, 47, 267], many efforts have been made to mimic the Outer Plexiform Layer (OPL) circuitry in silicon retinas [171, 172, 174, 175, 183, 265, 268, 269], or to implement the model in software for image processing [179, 270–272]. In the latter case, the normalization giving the adaptation is implemented numerically after digitalization which gives rise to noise amplification, especially in dark scenes. Consequently, considerable interest has focused during the last twenty years on developing a silicon retina giving an OPL-like response over the entire visible spectrum, also thanks to the latest advances in retinal implant systems [273, 274].

The first example of an auto-adaptive silicon retina was presented in [171], where a logarithmic photoreceptor was used to handle transient changes in light in a 1-decade range, while light adaptation within a 1-decade range was obtained by "subtracting" a local spatio-temporal average. This circuit improved the contrast resolution of equally illuminated areas in comparison with standard logarithmic photoreceptor retinas, but there was no improvement in the low signal-to-noise ratio inherent to the logarithmic amplification. In [172], a modified version of this chip was compared with the OPL response described in *Necturus* [173] (see Section 3.2.3), showing light adaptation in a 5-decade range but sensitivity to luminous changes within a range of only 0.5 decades.

To overcome these limitations, a more biologically inspired solution was subsequently developed in a study by [174], which consisted in modulating the synaptic strengths locally to control the sensitivity and including cone-to-cone gap junctions to attenuate the noise. Although the sensitivity was improved in this way from 0.5 to 2 decades, the adaptation to light was not satisfactory because of the circuit deviations resulting from the increasing inter-receptor coupling strength. A good compromise between contrast sensitivity and light adaptation was reached in [175], which gave light adaptation in a 6-decade range and sensitivity in a 1-decade range. However, the steady-state response of this pixel was found to increase with the light intensity (i.e., the photodiode current) and the transient response was not always monotonic when large lighting variations occurred (see Fig. 2.13 in [176]). The Delbrück adaptive pixel was also found in studies on optic flow measurements to have little practical use in situations where changes in the light greater than 1 decade are liable to occur (see Figs. 7(b),7(d),7(j) and 7(l) from [177]). In [178], the Gamma correction method presented in [179] for local tone-mapping purposes

was improved by digitally normalizing the pixel output directly in VLSI in line with the Michaelis-Menten law: only a few preliminary results on light adaptation and contrast sensitivity were presented, however, in that study.

Other solutions not involving the use of auto-adaptive elements have been suggested. In [265], a subretinal stimulator was endowed with light adaptation by shifting the dynamic input range via an externally generated signal. The results presented showed that photoreceptor adaptation was achieved in a 7-decade range. In [269], a silicon retina was provided with a wide dynamic operating range and a high contrast sensitivity by applying a spatial and temporal filtering process based on resistive networks. The results obtained showed that the pixels could deal with 4 decades of luminosity changes, but their contrast responses depended on the external voltage controlling a reset transistor.

As far as we know, no artificial retinas have ever been endowed up to now with pixels with the following features at once: (i) auto-adaptation to the mean local luminosity over a range as wide as the visible spectrum; (ii) constant sensitivity to luminous changes, i.e. contrasts, at any average luminosity in the operating range; (iii) reliable response even in the presence of sudden large changes in the luminosity (i.e., without causing circuit saturations or deviations).

In this paper, it is proposed to present: (i) a novel analog silicon retina featuring auto-adaptive pixels that obey the Michaelis-Menten law faithfully in a 7-decade range without causing any saturation of the VLSI transistors, while keeping an effective resolution of the integrated analog-to-digital conversion of about 9 bits; (ii) a method of characterizing silicon retinas, which can be used to accurately assess the pixels' response to transient luminous changes within a ± 3-decade range and to changes in the steady-state intensity within a 7-decade range. We have called this novel pixel the M^{2}APix, which stands for the Michaelis-Menten Auto-adaptive Pixel.

The present artificial retina consists of a $2 \times 2\,mm$ CMOS circuit comprising four lines of six auto-adaptive pixels, and a digital interface giving a fast serial read-out of up to $1\,MHz$ connecting the retina directly to an external microprocessor or microcontroller. The adaptation time constant of the M^{2}APix can be changed by means of an external capacitor, providing additional flexibility to eventually meet the application's requirements in term of preferred bandwidth.

The biological background to this study is presented in Section 3.2.3. The chip implementation is presented in Section 3.2.4, and a detailed description of our auto-adaptive pixel is provided in Section 3.2.5. The method of characterization used is presented in Section 3.2.8, and the results obtained using this method are presented and discussed in Section 3.2.9. A comparison between the M^{2}APix and the Delbrück pixel present on the same silicon retina is proposed in Section 3.2.10. Some conclusions are reached in the last section.

3.2.3 Biological background

Light adaptation of the photoreceptors present in human and animal retinas has been extensively studied in a large number of species since the early 50's, using both intra and extracellular methods [36–43]. In all

these studies, the relationship between light stimuli and photoreceptor responses has been documented, both in the dark and with background illumination, via an adaptation process described by the so-called Michaelis-Menten equation [44]:

$$V = V_m \frac{I^n}{I^n + \sigma^n}, \tag{3.1}$$

where V stands for the photoreceptor's response and V_m is its maximum value; I denotes the light intensity and n usually ranges from 0.7 to 1; σ is the adaptation parameter, corresponding to the light intensity giving half of the maximum response.

The first micro-electrode recordings of rod and cone responses were obtained on saltwater fish (*Gerridae*) by Svaetichin in 1953 [45]. In his pioneering study, Svaetichin discovered the S-potentials, as they were subsequently called by Oikawa *et al.* in [46], which stands for "slow potentials", referring to the slow adaptation process which occurs in the photoreceptor potentials when they are exposed to flash lights against a steady background.

However, the first mathematical description of the cone response given by equation (3.1) was provided by Naka and Rushton in the case of the freshwater fish (*Cyprinidae*) [36]. The equation (3.1) where $n = 1$ is therefore also known as the Naka-Rushton law. The same model was subsequently validated and applied to turtles' cones by Baylor *et al.* [37] and to monkeys' cones by Boynton and Whitten [38], who introduced the exponent $n < 1$ for the first time. Many studies were then carried out on vertebrates and invertebrates, all confirming the equation in (3.1) with various values of n and sometimes with different interpretations of the adaptation parameter σ (in the salamander [39], gecko [40], frog [41], locust, fly and dragonfly [42], for instance, and in the human fovea [43]).

Figure 3.1 shows the responses of dark- and light-adapted red cone photoreceptors recorded intracellularly in the retina of the turtle (*Pseduemys Scripta Elegans*) by Normann and Perlman [47]. As can be seen from this figure, the function $V(I)$ defined in (3.1) gives rise in the $Log(I)$ domain to curves with a fairly smooth "S" shape (continuous curves), where the slope of the "S" is given by the value of n ($n = 1$ in that case) and the lateral shift by the value of σ.

Based on the S-shaped curves shown in Fig. 3.1, two main features of the light-adaptation behavior can be described by the incident-light model [50]:

- as the background lighting changes, the entire S-shaped curve shifts along the light intensity axis, which corresponds to a change in the sensitivity of the photoreceptor in the neighborhood of the background light. In fact, after reaching a peak value caused by an increase/decrease in the intensity of the light (data points), the potential V gradually returns to a steady-state value, reflecting its adaptation to the background. This decrease/increase in V corresponds to a "slow" increase/decrease in the parameter σ (see equation (3.1));

- as the background illumination increases, the operating point of the photoreceptor increases correspondingly (small horizontal lines in Fig. 3.1), which means that because of the non-linearity of the curve, the response to a given increment/decrement in the stimulus becomes smaller/larger at higher background levels. This process known as "response compression" was introduced for

the first time by Boynton and Whitten in [38]. The slope of the curve around the operating point defines the contrast sensitivity.

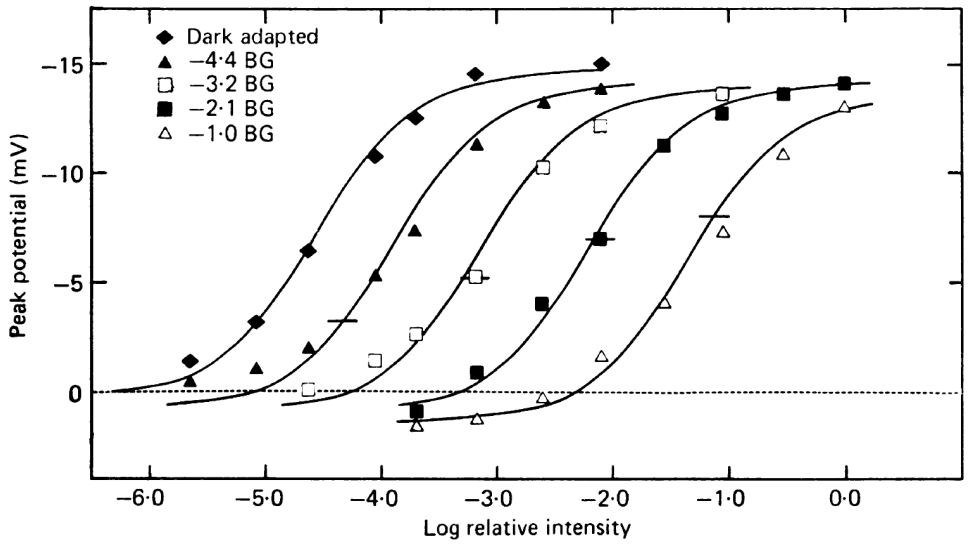

Figure 3.1: S-shaped curves corresponding to dark- and light-adapted response curves recorded in a red cone of the turtle. The peak of either the incremental or decremental response measured from the dark-adapted potential recorded before the background onset (dashed line) is plotted as a function of the log of the test pulse intensity which elicited each response. The steady hyperpolarization produced by each background lighting condition is given by the intersection between the intensity-response curve and the small horizontal line. The continuous curves were drawn from a single template which describes the function $V = V_m \frac{I}{I+\sigma}$. Adapted from [47].

3.2.4 Chip implementation

In this section, we present our 2-D photosensor array, featuring a silicon retina composed of 24 auto-adaptive pixels of two different kinds. A picture of the chip package, with a zoom on the retina, is presented in Fig. 3.2.

The retina consists of a $2 \times 2\,mm$ CMOS circuit designed using the $350\,nm$ XFAB standard CMOS process. To facilitate the integration of the chip into custom-made Printed Circuit Boards (PCBs), the circuit was encapsulated in a standard $9 \times 9\,mm$ (LCC24) package with 24 pins. Four rows of auto-adaptive pixels with a $254\,\mu m$ diameter N-well/P-substrate photodiode were implemented in the PCB, as shown in Fig. 3.2(a). The photodiode is mostly sensitive to red light ($\lambda \approx 650\,nm$) and its sensitivity S_{ph} is equal to $1.1 \times 10^{-8} \frac{Am^2}{W}$. The retina is completely self-biased and can be read out by means of digital serial read-out architecture implemented directly on the chip. The main functional blocks implemented on the chip can therefore be summarized as follows:

- two rows of six Michaelis-Menten auto-adaptive pixels, that we called M^{2}APix (see Section 3.2.5

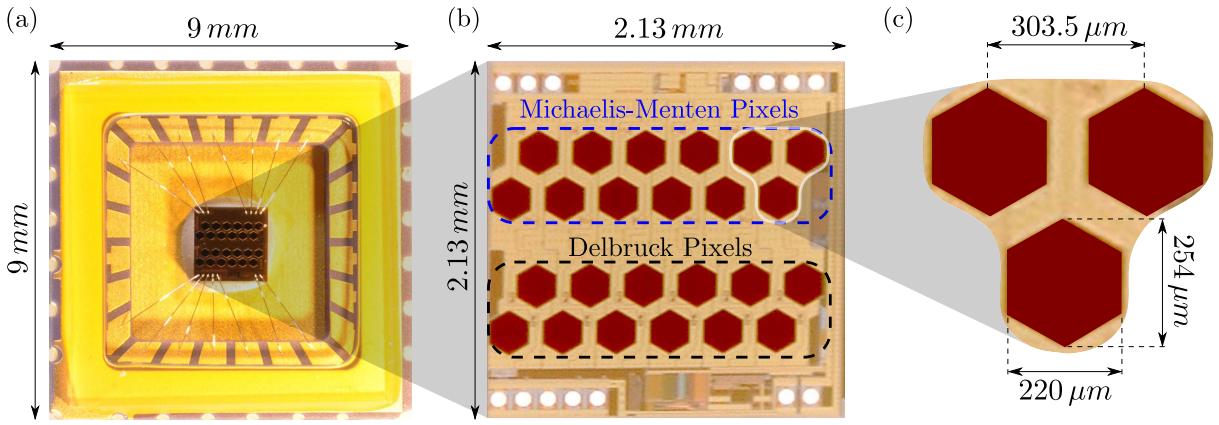

Figure 3.2: (a) The silicon retina in its $9 \times 9\,mm$ package; (b) Magnified view of the silicon retina composed of 12 Michaelis-Menten pixels presented in this study, and 12 additional Delbrück pixels; (c) Magnified view of 3 Michaelis-Menten pixels giving the photodiode's dimensions and the inter-receptor distance.

for details);

- a low-pass filtered current-averaging cell and the reference voltage, required by the Michaelis-Menten architecture (see Section 3.2.5);
- a bias generator providing the polarization currents required for the circuit to operate properly;
- a reference voltage for the analog-to-digital converter (ADC);
- a digital serial interface, which includes the digitizing of the pixel output signals via a 10-bit ADC, and a direct serial communication bus (see details below).

In addition, two rows of six pixels of the Delbrück type (see [175] for more details) were implemented on the same chip.

The photodiodes were aligned on two horizontally staggered rows so that the hexagons fit together like a puzzle, recalling the shape and the arrangement of insects' hexagonal ommatidia (Fig. 3.2). This pattern of alignment of the photodiodes, which is particularly suitable for detecting luminous contrasts in the main direction, also makes it possible to sense light variations in any other direction.

The analog signals originating from each pixel were digitized on-chip so that they could be directly processed by an external microprocessor or microcontroller, saving the power consumption and the computational resources of the latter for further data processing. Since the pixel type is selected by a digital input, only one set of pixels can be converted at a time. The twelve M²APix outputs are low-pass filtered with a cut-off frequency of $300\,Hz$ before being digitally converted, giving a minimum sampling frequency of $600\,Hz$ in order to prevent the occurrence of aliasing. An integrated DC reference voltage was connected to the ADC as an additional pixel output for testing and calibration purposes. To improve the LSB voltage, the dynamic input range of the ADC can be reduced by using an external voltage

source. A synchronous direct-connection protocol similar to that used in the artificial compound eye CurvACE [219] was adopted, as it provides a very simple, robust solution.

A functional diagram of the interface protocol is given in Fig. 3.3. For the serial communications, a similar internal state machine working with a maximum frequency of $1\,MHz$ to that adopted in the CurvACE sensor [219] was used to transfer the data to an external device.

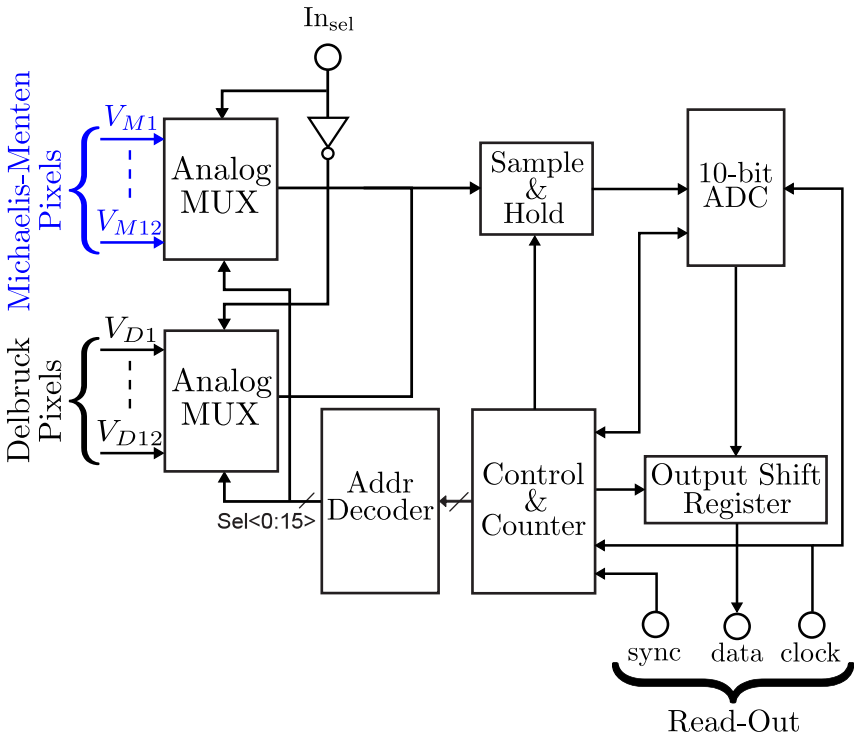

Figure 3.3: Block diagram of the retina's serial synchronous read-out interface (see [275]).

3.2.5 M^{2}APix: Michaelis-Menten auto-adaptive pixel

In this section, we present our novel auto-adaptive pixel implementing the Michaelis-Menten model in analog VLSI, as described in Section 3.2.3. The theoretical basis of the analog circuit is first presented, and an example of the auto-adaptive pixel's response is then given to illustrate the behavior of the M^{2}APix and show how the pixel is related to S-shaped curves such as those presented in Fig. 3.1.

3.2.6 Circuit description

The block scheme depicted in Fig. 3.4(a) gives an overview of the Michaelis-Menten pixel implementation. All the blocks in the area delimited by the dashed lines belong to a single pixel. These blocks are therefore replicated twelve times, whereas the two blocks outside the dashed-line area are common to all twelve pixels.

To implement the Michaelis-Menten function in (3.1), we adopted the current normalizer model presented in [276] (Chapter 6, pp. 148-150), with an arbitrary number of current inputs, and patented

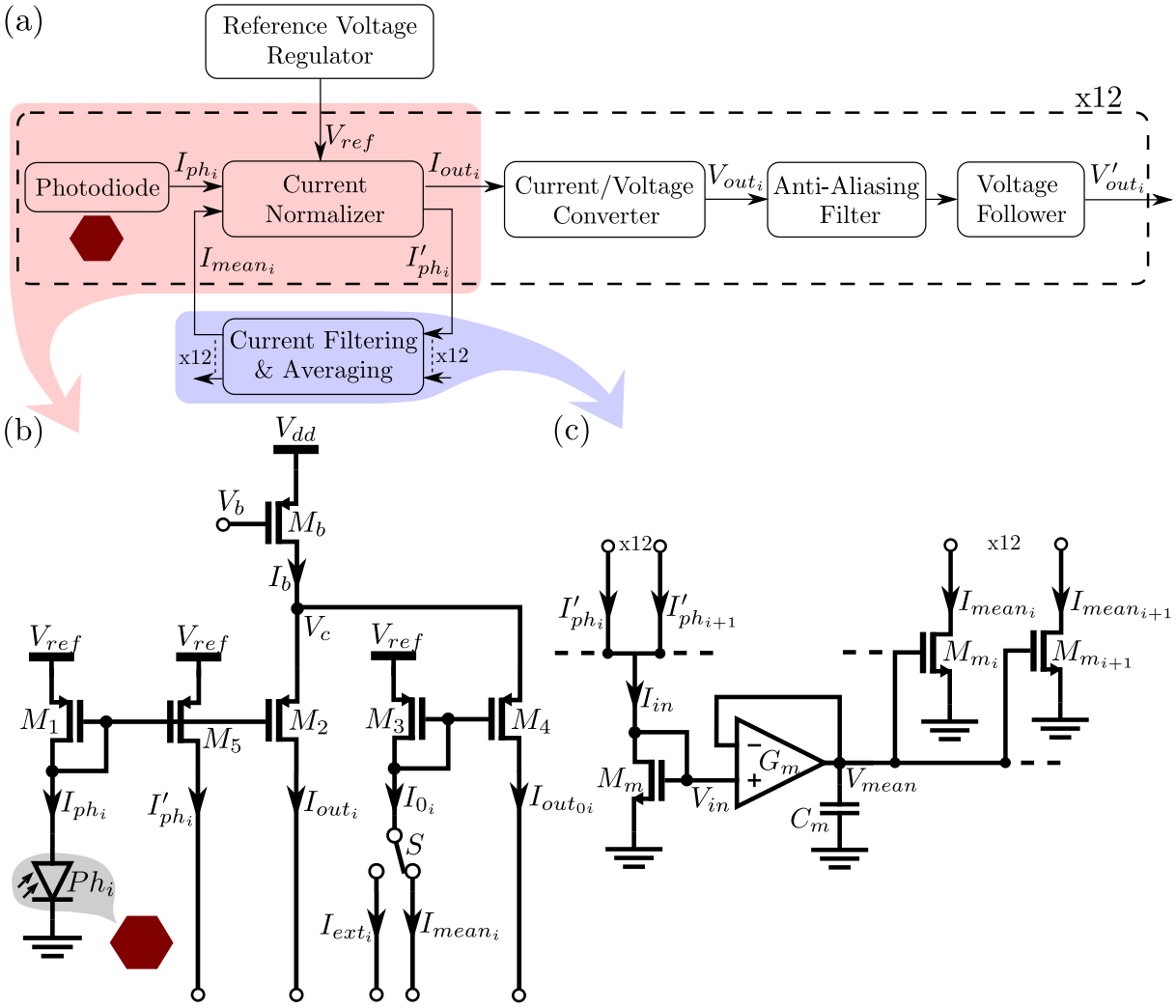

Figure 3.4: (a) Block diagram of the M²APix: Michaelis-Menten Auto-adaptive Pixel. The blocks in the dashed-line area, which are replicated 12 times, belong to a single pixel, whereas the two blocks outside the dashed-line area are common to all twelve pixels. (b) Hardware implementation in VLSI of an elementary auto-adaptive pixel (photodiode and current normalizer), the output signal of which is noted I_{out_i}. A switch S can be used to select either I_{0_i} as the mean current I_{mean_i} provided by the built-in averaging circuit, or an external current I_{ext_i} provided by an external circuit. (c) Hardware implementation in VLSI of the filtering and averaging circuit computing the mean current of the 12 mirrored photodiode currents (I'_{ph_i}) produced by the 12 normalizer circuits.

in [277] in the context of photosensing. In the present application, two current inputs are required: one for the photodiode current I_{ph_i} and one for the current I_{0_i} corresponding to the average illuminance. The functional scheme of the current normalizer is presented in Fig. 3.4(b). A switch S can be used to select either I_{0_i} as the mean current I_{mean_i} provided by the built-in averaging circuit or an external current I_{ext_i} provided by an external circuit. In what follows, we take $I_{0_i} = I_{mean_i}$.

The scheme adopted here improves the functioning of the current normalizer by adjusting the V_{ref} voltage to a different value from that of V_{dd} $(3.3\,V)$. The auto-adaptive pixels we designed work efficiently in a wide range of luminosities corresponding to a photodiode current ranging from about $20\,pA$ to $20\,\mu A$. The V_{ref} voltage optimizes the functioning of the system at low currents by preventing the current source M_b transistor from saturating.

Operating principle of the normalizer

If the transistors $M_1, \ldots, M_4$ have the same dimensions and are working in their sub-threshold region, the current output I_{out_i} can be expressed as follows:

$$I_{out_i} = I_b \frac{I_{ph_i}}{I_{ph_i} + I_{mean_i}}, \tag{3.2}$$

where $I_b = 50\,nA$ and the index $i = 1 \ldots 12$ gives the number of pixels. For the sake of simplicity, we will drop the index in what follows.

To obtain the same auto-adaptation to light as that which occurs in animals' eyes, I_{mean} has to be a representative value of the background luminosity perceived by the artificial retina (see Section 3.2.2 and 3.2.3). Accordingly, the computation of I_{mean} must reflect only the low-frequency changes in the light perceived by all the photoreceptors in the retina. The current I_{mean} is therefore the average value of copies of all 12 photodiode currents (I'_{ph_i}) filtered with a first order low-pass filter. As shown in Fig. 3.4(c), the low-pass filter is provided by a gm-C structure using an operational transconductance amplifier (OTA) G_m and an external capacitor C_m. In particular, if the external capacity is set at $100\,nF$, as done in our tests, the cut-off frequency will be $150\,mHz$.

As shown in Fig. 3.4(a), the output current of the normalizer is converted into a voltage via a high gain transimpedance amplifier (TIA). This high-factor current-to-voltage conversion R_f is obtained using a low-transconductance OTA in the feedback loop of an operational amplifier stage. The output voltage V_{out} can therefore be expressed as follows:

$$V_{out} = -R_f I_{out}, \tag{3.3}$$

where R_f is set at $17.5\,M\Omega$ via the OTA transconductance.

To prevent the occurrence of aliasing due to the sampling frequency of the digital conversion, a first-order low-pass filter with a gm-C structure is added to each pixel. A cut-off frequency of $300\,Hz$ is achieved by means of a low-transconductance OTA and an internal capacitance.

Lastly, a voltage follower helps the pixel to drive the input sampling capacitance of the ADC.

The output pixel voltage can then be written as follows:

$$V_{out} = -R_f I_b \frac{I_{ph}}{I_{ph} + I_{mean}} + V_{BG}, \tag{3.4}$$

where $V_{BG} \approx 2.3\,V$ denotes the band-gap voltage due to the intrinsic functioning of the various stages.

3.2.7 M^{2}APix response

The fact that the term I_{mean} in equation (3.4) corresponds to the average luminosity constitutes a key point in the adaptive behavior of the pixel.

Let us assume that in the absence of any optical lenses placed on the retina, all the pixels are exposed to the same light intensity. If no changes or only very slow changes in the luminosity occur, all the photodiode currents and their average will be identical. Therefore by substituting $I_{mean} = I_{ph}$ into (3.4), the steady-state value of the pixel's output can be obtained:

$$V_{out_0} = -\frac{R_f I_b}{2} + V_{BG}, \tag{3.5}$$

which is a constant value depending only on the operating current I_b and not on the photodiode currents.

This feature is the main difference with respect to the biological findings, which on the contrary, show the existence of a logarithmic increase in the steady-state response with respect to the luminosity, giving rise to the so called "response compression" (see Section 3.2.3). In fact, while the contrast sensitivity of the OPL varies, depending on the average luminosity (see for instance the slope around the small horizontal line of the full-triangle curve in comparison with the empty-triangle one in Fig. 3.1), our silicon retina shows the same contrast sensitivity whatever the average luminosity. In other words, an object showing a contrast of 10% will generate the same signal amplitude at the output of the M^{2}APix under both low and high luminosity levels.

In order to explain this behavior more clearly and show how it is related to S-shaped curves such as those presented in Fig. 3.1, an example of the pixel's response was plotted as shown in Fig. 3.5(a).

At the time t_0, all the photodiode currents and the pixel's output are stabilized at I_{ph_0} and V_{out_0}, respectively. Then, after the time T_s, a step change in the luminosity occurs in front of the silicon retina, making all the photodiode currents "rapidly" change from I_{ph_0} to I_{ph_i} (red line), while the average current changes much more slowly (magenta line). During the stabilization of the average current, the pixel's output first increases from V_{out_0} to the peak value V_{out_i}, and then decreases until it reaches its stable baseline value V_{out_0} (blue line) again. We take *rise time* (T_r) to denote the time taken by V_{out} to go from V_{out_0} to 90% of $(V_{out_i} - V_{out_0})$ and *fall time* (T_f) the time taken by the signal to go from V_{out_i} to 90% of $(V_{out_0} - V_{out_i})$. It can be shown that T_r and T_f mainly depend on the photodiode and transistors time-constant and the averaging block time-constant ($\tau_m = \frac{C_m}{G_m}$), respectively.

As $T_r \ll \tau_m$, it can be assumed that when $V_{out} = V_{out_i}$, I_{mean} does not change ($I_{mean} \approx I_{ph_0}$) independently from I_{ph_i}. Therefore, the points defining the S-shaped curve correspond to the peak values

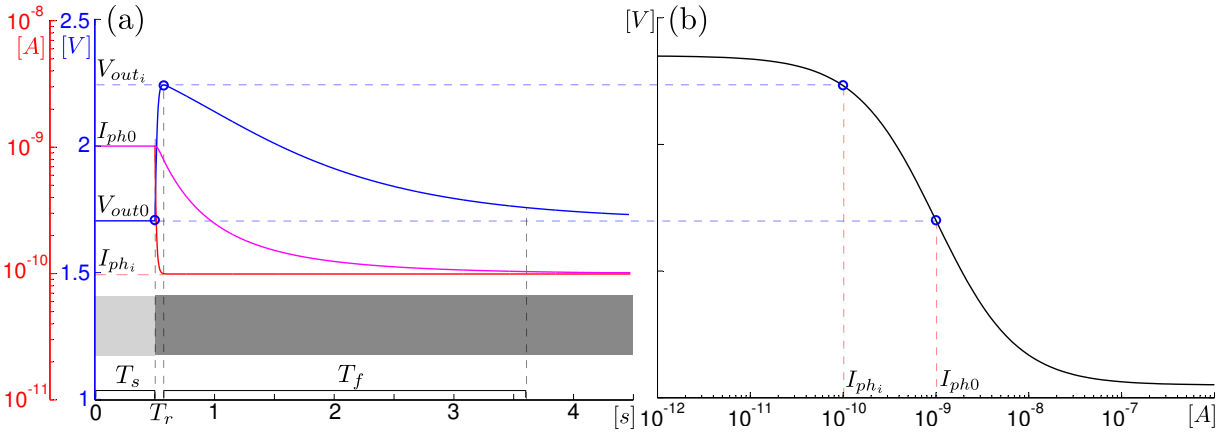

Figure 3.5: (a) Simulated example of the adaptive pixel's response (blue line) to a step change in the light intensity (light to dark gray stripe). The red and magenta lines stand for the photodiode current and the average current, respectively. (b) Theoretical S-shaped curve based on equation (3.4), giving the peak values of the pixel's response V_i to step changes in the photodiode current from I_{ph_0} to I_{ph_i}.

V_{out_i} reached at all the step values I_{ph_i} after the same initial value I_{ph_0} (Fig. 3.5(b)). It is worth noting that when the value of I_{ph_i} reaches up to ± 1 decades of I_{ph_0}, the pixel shows a logarithmic sensitivity to changes in the lighting conditions (see the sloping part of the curve around the operating point), while the sensitivity decreases drastically in response to greater changes in the light.

Thanks to the intrinsic properties of the normalization, the peak values depend only on the ratio $\frac{I_{ph_i}}{I_{ph_0}}$, resulting in a horizontal shift of the S-shaped curve depending on I_{ph_i} in the same way as the curves in Fig. 3.1.

In addition, as I_{ph} is linearly proportional to the luminous intensity, we can assume the presence of a luminous contrast defined by the Michelson formula: $c_i = \frac{I_{ph_i} - I_{ph_0}}{I_{ph_i} + I_{ph_0}}$. Substituting the inverse of this formula into equation (3.4), i.e. $I_{ph_i} = \frac{1+c_i}{1-c_i} I_{ph_0}$, we obtain:

$$V_{out_i} = -R_f I_b \frac{c_i + 1}{2} + V_{BG}. \tag{3.6}$$

The pixel can therefore be said to give a linear contrast response, and the contrast resolution is given by the coefficient of c_i in equation (3.6).

Lastly, an AC noise simulation was performed with a white noise at the input (corresponding to the shot noise of the photodiode), to obtain the Root Mean Square (RMS) of the output noise and consequently the minimum detectable contrast for different values of the average luminosity, i.e. the DC photodiode current. The RMS values are given by integrating the output noise in $[10^{-5}, 10^{8}]\, Hz$. The minimum detectable contrast can be defined as the contrast that gives rise to a transient response of the output signal ± 6-fold the RMS noise.

Figure 3.6 shows the RMS of the simulated output noise (blue) and the minimum detectable contrast (red) with respect to the photodiode current. We can notice that the noise is decreasing with the DC

photodiode current. At very low background luminosity, the noise is dominated by the input white noise of the photodiode which is amplified by the gain of the circuit. At higher luminosity, this gain is smaller, so the simulated input noise becomes negligible and the output noise is nearly equal to the transistors' noise. In any case, the RMS values obtained are always very low compared to the output variations $(V_{out_i} - V_{out_0})$ in the transient response, as shown in the simulated response in Fig. 3.5. Consequently, the minimum detectable contrast is very low over the entire operating range of the average luminosity, varying from $\pm 1.1\,\%$ at $1\,Lux$ to $\pm 0.4\,\%$ at $10^5\,Lux$.

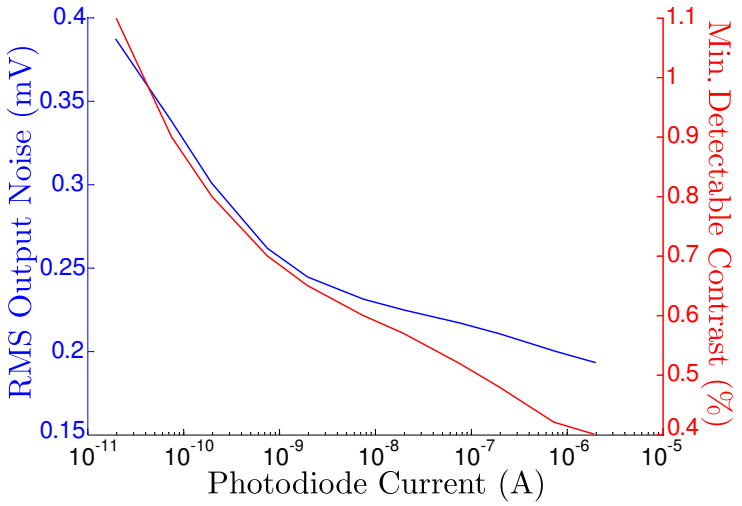

Figure 3.6: Root Mean Square (RMS) of the simulated output noise (blue) and corresponding minimum detectable contrast (red) with respect to the photodiode current. The RMS values are given by integrating in $[10^{-5}, 10^{8}]\,Hz$ the output noise obtained with an AC noise simulation which takes into account all the transistor noises and a white noise for the input photodiode. The minimum detectable contrast can be defined as the contrast that gives rise to a transient response of the output signal ± 6-fold the RMS noise.

3.2.8 Method of characterization

In the studies presented in Section 3.2.2, the light adaptation and contrast sensitivity of silicon retinas were often tested by applying a series of lighting steps (AC light) in addition to various background lights (DC light). As a result, the pixels' responses have often been described in terms of the stimulus intensity, and direct comparisons can therefore be made with the biological findings (see Section 3.2.3). However, the method used to characterize pixels' responses is sometimes not clear or has not even been described at all, which makes comparisons with other results very difficult.

A standard method is presented here for accurately characterizing pixels' responses to luminous changes of up to ± 3 decades in a 7-decade mean luminosity range by implementing a single light source, which has been called the *Lighting Box* (Fig. 3.7).

The Lighting Box consists of a $50 \times 25 \times 25\,mm$ 3-D printed box with a $10 \times 10\,mm$ aperture. The Lighting Box also includes a Printed Circuit Board (PCB) which accurately controls the light intensity

of a red Light Emitting Diode (LED) (TLWR7600, Vishay Semiconductors) by means of a specific digital current driver (ADN8810, Analog Devices). The PCB with the LED and the PCB supporting the retina to be characterized are fixed to each side of the box so that both the LED and the silicon retina can fit into the aperture, facing each other inside the box (see Fig. 3.7(b)). The box also contains an optical filter support, which can be inserted between the LED and the retina in order to attenuate the LED's intensity and thus characterize the pixel's response in the illuminance range of interest.

The following main tools were used for this purpose:

- The Lighting Box described above.
- A NI Single-Board RIO-9683 Acquisition Device provided by National Instruments. The board features a $400\,MHz$ real-time processor with 128 MB DRAM and includes an integrated real-time controller and a 2-Million-Gate reconfigurable FPGA programmed using LabVIEW software including Real-Time and FPGA modules.
- An Explorer 16 development board provided by Microchip, which includes a dsPIC 33FJ128GP804 micro-controller working at a sampling rate of $2\,kHz$. This device was programmed using Matlab/Simulink with a toolbox specifically developed for use with Microchip dsPIC micro-controllers.

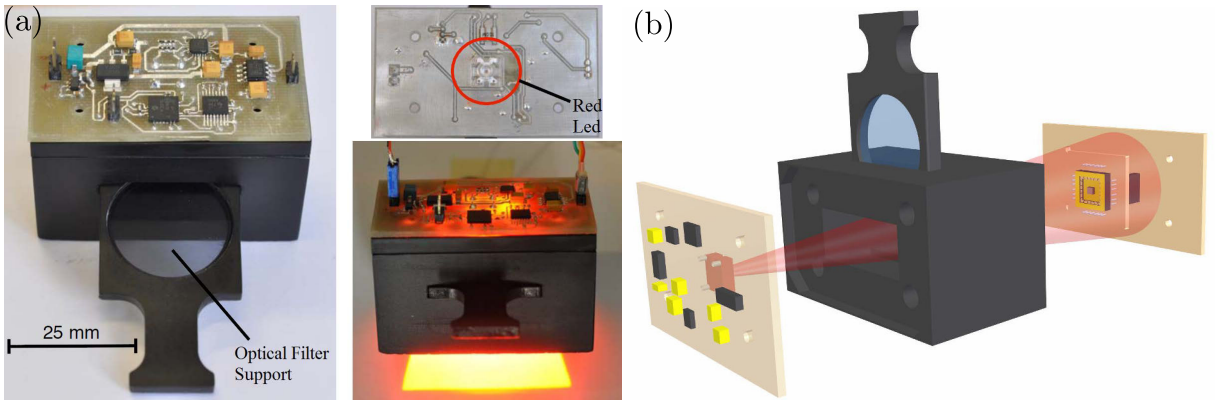

Figure 3.7: (a) Pictures and (b) exploded view of the *Lighting Box* composed of a PCB with a red LED ($\lambda \approx 618\,nm$) and an optical filter support. The direct control of the LED current makes the illuminance vary in a 3-decade range. Additional optical filters (neutral density filters) were used to drastically increase the mean luminosity range from 3 to 7 decades.

The block diagram in Fig. 3.8 gives an overview of the hardware setup and communication flow involved in the method of characterization.

Characterization procedure

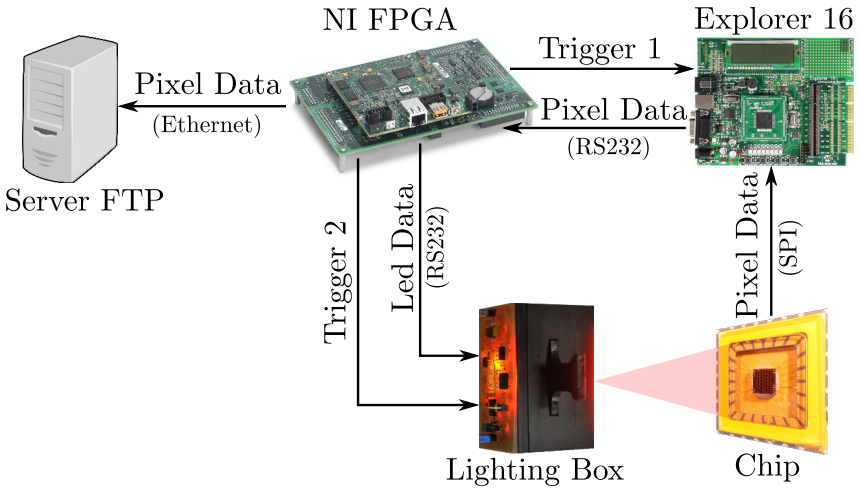

Figure 3.8: Block diagram of the hardware setup and communication flow involved in the pixel characterization procedure.

The 12 pixels' output signals were acquired while they were being exposed to step changes in the luminous intensity, as described in Section 3.2.7. During the overall acquisition process, the FPGA acts as the master component handling and synchronizing the communications with the Lighting Box in order to drive the intensity of the LED, and with the Explorer 16 in order to acquire data from the chip. Based on Fig. 3.5 and Fig. 3.8, the i-th step in the procedure can be described as follows:

1. The FPGA sends the Lighting Box a packet of four bytes (Led Data) containing the information about the initial value (I_{LED_0}) and the step value (I_{LED_i}) of the LED intensity.

2. As soon as the Lighting Box receives the packet, it sets the LED intensity at the initial value I_{LED_0}.

3. After waiting for a time (T_{wait}), which is usually the time required for the pixel output to reach its steady-state value, the FPGA sends Explorer 16 a trigger signal (Trigger 1) making it start sending the data acquired. The steady-state value of the pixels is thus acquired before the lighting change occurs.

4. A short instant later (T_s), the FPGA sends the Lighting Box a second trigger signal (Trigger 2) making it switch the LED intensity from I_{LED_0} to I_{LED_i} (see Fig. 3.5).

5. The Explorer 16 sends the FPGA the appropriate number of samples (Pixel Data), depending on the sampling frequency.

6. The FPGA stores the data acquired in a FTP server and goes back to step 1 to deal with the next pair I_{LED_0}, I_{LED_i}.

The current-irradiance characteristic of the LED was assessed by measuring the LED's irradiance with a radiometer (ILT1700, International Light Technologies). To obtain a good idea of what the photodiodes perceive, the radiometer was placed in front of the LED at the same distance as the chip. Therefore,

without any loss of generality, I_{LED} can be taken to stand for the LED's irradiance instead of the LED current. Since the photodiode current is linearly proportional to the irradiance via the sensitivity S_{ph}, as defined in Section 3.2.4, equation (3.4) does not have to be changed even if we take V_{out} to be a function of I_{LED}.

As we were interested in characterizing our auto-adaptive pixel in 7 decades of photodiode current (see Section 3.2.6) using a LED covering three decades, four neutral optical filters (NG filters, Schott) were used: the 1 and $2\,mm$ NG3 type for dealing with 1- and 2-decade attenuation, respectively, and the 2 and $3\,mm$ NG9 type for dealing with 3- and 4-decade attenuation, respectively. To check the full set of S-shaped curves within ± 3 decades about I_{ph_0} (see Fig. 3.5(b), for example), a complete pixel characterization was carried out by merging the data obtained with the various optical filters. In particular, as the latter give contiguous 1-decade attenuation, S-shaped curves were obtained by merging and averaging the peak values V_{out_i} acquired at the same initial irradiance with several filters. For instance, the S-shaped curve centered in $0.1\,\frac{W}{m^2}$ (Fig. 3.5(b)) was obtained by merging the peak values obtained with $I_{LED_0} = 0.1\,mA$ without any filter, $I_{LED_0} = 1\,mA$ with a 1-decade attenuation filter, $I_{LED_0} = 10\,mA$ with a 2-decade attenuation filter and $I_{LED_0} = 100\,mA$ with a 3-decade attenuation filter.

3.2.9 M^{2}APix characterization results

In this section, it is proposed to present and discuss the experimental data obtained when our auto-adaptive silicon retina was exposed to changes in the light, as described in Section 3.2.8.

To characterize the pixel's response in the light-adapted condition, the LED light changes were triggered when the pixel's output signal reached its steady-state value. In other words, by referring to the steps in the characterization procedure presented in Section 3.2.8, the waiting time T_{wait} in Step 3 was equal to $5T_f$, where T_f denotes the fall time, as described in Section 3.2.7.

M^{2}APix transient response

As mentioned in Section 3.2.7, we were interested in measuring the peak values V_{out_i} of the twelve pixels with each pair of photodiode currents I_{ph_0}, I_{ph_i} corresponding to a step change in the light intensity (Fig. 3.5). Pixel output signals were recorded with various pairs of LED irradiance values I_{LED_0}, I_{LED_i} in the $[10^{-5}, 10^2]\,\frac{W}{m^2}$ range at a sampling frequency $f_s = 2\,kHz$, in order to accurately determine the instant at which the peak value occurred, using the method described in Section 3.2.8. The set of all the mean values obtained when $I_{LED_i} = I_{LED_0}$ corresponds to the steady-state response of the pixel and constitutes an important auto-adaptive characteristic of the M^{2}APix, which will be discussed below.

Some examples of the pixel's response with $I_{LED_0} = 1\,\frac{W}{m^2}$ are presented in Fig. 3.9.

Figure 3.9 shows the auto-adaptation of the pixel's time response. According to the simulated example presented in Section 3.2.7, the pixel's output rapidly increases from the steady value V_{out_0} to a peak value V_{out_i} and then returns slowly to the steady value regardless of the contrast, which was obtained

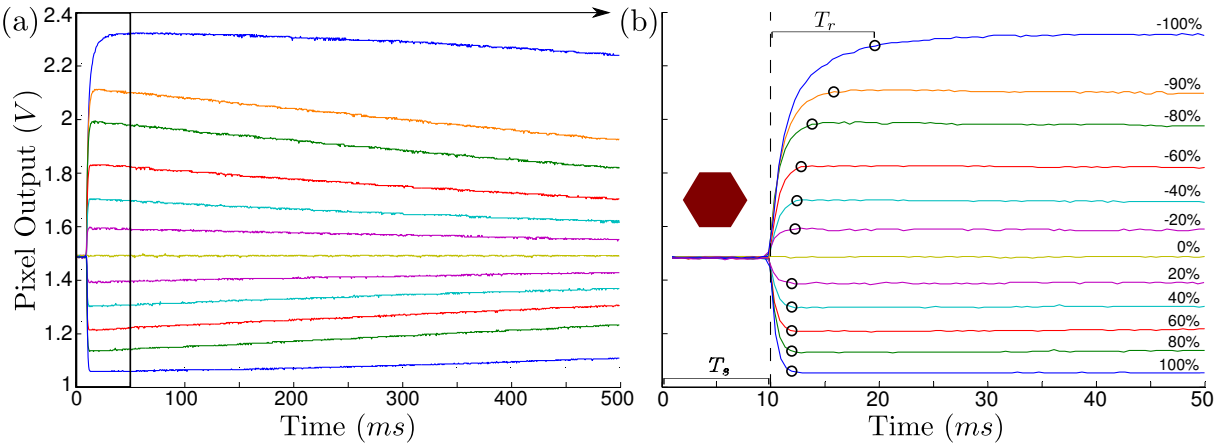

Figure 3.9: (a) Example of M²APix responses to step changes in the LED irradiance (I_{LED_i}), starting with the same initial irradiance ($I_{LED_0} = 1\,\frac{W}{m^2}$). (b) Zoom of the temporal pixel responses shown in Fig. 3.9(a) ranging between 0 and $50\,ms$. To be able to distinguish more clearly between the steady-state responses and the transient responses, the step changes were delayed by $10\,ms$ (T_s) after the acquisition procedure had started. The black circles amount to 90% of the peak values V_{out_i} and give qualitative information about the rise time (T_r). The contrast values are given by the Michelson formula $c_i = \frac{I_{LED_i} - I_{LED_0}}{I_{LED_i} + I_{LED_0}}$.

here by making step changes of I_{LED_i}. However, as shown in Fig. 3.9(b), the rise time (T_r), namely the time required to reach 90% of the peak value (black circles), is nearly constant with positive contrasts ($I_{LED_i} > I_{LED_0}$), but depends on the contrast with negative ones ($I_{LED_i} < I_{LED_0}$).

Figure 3.10 shows the mean rise time over the 12 pixels plotted with respect to the contrast. Each point corresponds to the time (T_r) required to reach 90% of the peak value V_{out_i}, as depicted in Fig. 3.9(b), for every step change (I_{LED_0}, I_{LED_i}).

By looking at Fig. 3.10, we can notice that the rise time depends on both the contrast and the average luminosity. This coupling between contrast and rise time is directly due to the current-mode functioning of the system. A high negative contrast corresponds to a normalizer's output current tending to zero, as we can see by considering $I_{ph_i} \ll I_{mean_i}$ in equation (3.2). The time constants of the output transistors will therefore be higher and the output signal will be slower in this case. Furthermore, the time constant of the photodiode and the transistors increases for low current, i.e. low average luminosity, explaining the higher rise times for the pink and red data points. It is worth noting that the mean rise time for positive contrasts at medium/high average luminosity (blue and green data points) is about $1\,ms$ because of the anti-aliasing low-pass filter ($F_c = 300\,Hz$, see Section 3.2.6).

The rise time (T_r) and the fall time (T_f) depends on the bandwidth of the output signal that is determined mainly by the photodiode and transistors time constant and the current-averaging block time constant (τ_m). Such a bandwidth can be modified by modifying the external capacitor C_m, since $\tau_m = \frac{C_m}{G_m}$ (see Section 3.2.6). As the fall time T_f determines the time the pixel's output takes to reach

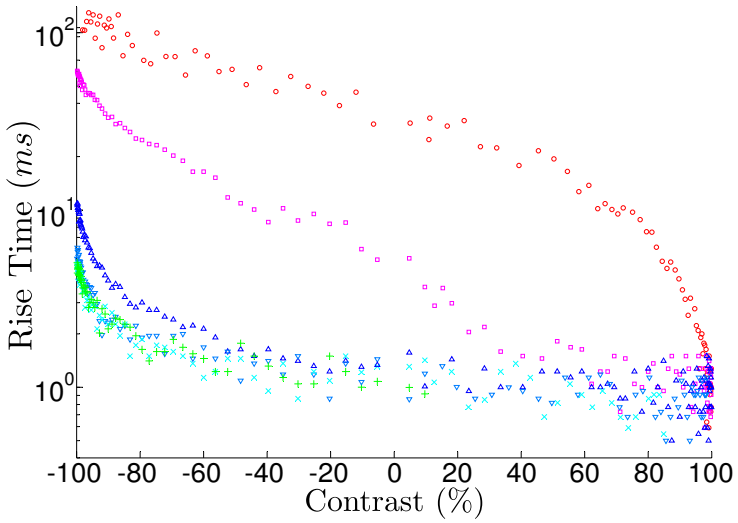

Figure 3.10: Average rise time with respect to luminous contrast. Each point corresponds to the time (T_r) required to reach 90% of the peak value V_{out_i}, as depicted in Fig. 3.9(b). Each color refers to a different initial irradiance value I_{LED_0}: red about 0.001 $\frac{W}{m^2}$, pink about 0.01 $\frac{W}{m^2}$, dark blue about 0.1 $\frac{W}{m^2}$, light blue about 1 $\frac{W}{m^2}$, cyan about 10 $\frac{W}{m^2}$, green about 100 $\frac{W}{m^2}$.

the maximum of its contrast sensitivity, i.e. the sensitivity at steady-state, it might be useful to set T_f as small as possible. If we consider an irradiance higher than $0.1\frac{W}{m^2}$ (about 100 Lux), the fall time T_f can be reasonably reduced up to $0.1\,s$ by changing the value of the external capacitor C_m to $5\,nF$, because the rise time T_r is lower than $0.01\,s$ for any contrast in this irradiance range. In fact, T_f should be about one decade greater than T_r to guarantee the correct functioning of the pixel in this luminosity range.

M^{2}APix S-shaped and steady-state response

The peak values obtained in response to step changes (I_{LED_i}) starting with several initial irradiances (I_{LED_0}) are presented in Fig. 3.11, with respect to (a) the LED irradiance and (b) the Michelson contrast c_i, defined as follows $c_i = \frac{I_{LED_i} - I_{LED_0}}{I_{LED_i} + I_{LED_0}}$.

By comparing the S-shaped curves in Fig. 3.11(a) with those in Fig. 3.1, it can be seen in the first place that our silicon retina shows in qualitative terms the same adaptation process as that observed in the OPL. In particular, the pixels auto-adapt to the average light in a 7-decade range while keeping a sensitivity of nearly $600\,\frac{mV}{Log(I)}$ in a range of about 2 decades (corresponding to the linear part of the curves on the log scale). Each S-shaped curve obtained was well defined within a range of 6 decades, which means that the circuit did not deviate, but remained consistent with the model, even when the light changed suddenly by anything up to ± 3 decades, which would correspond, for instance, to shifting from a very dark overcast sky to direct sunlight (see the lower half of the dark-blue curve). In addition, in line with equation (3.5), the steady-state response (black points) was almost constant throughout the 7 decades, giving the same contrast sensitivity whatever the average luminosity. Lastly, upon setting the high reference of the ADC at $2.4\,V$ (see Section 3.2.4 for further details about the integrated ADC), we

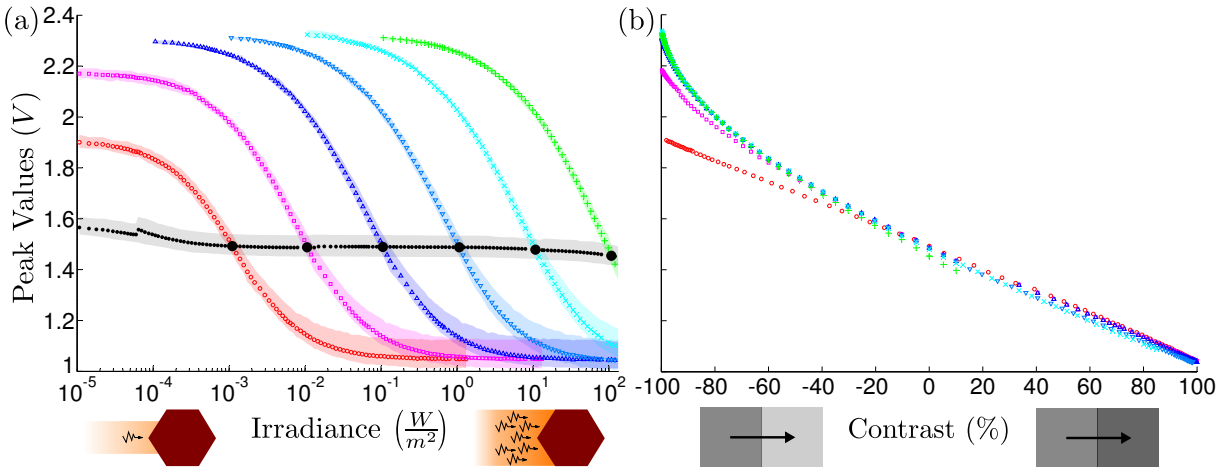

Figure 3.11: (a) S-shaped curves and steady-state responses of the 12 pixels to LED irradiance changes in a 7-decade range. Each color refers to a different initial irradiance value I_{LED_0} (red about 0.001 $\frac{W}{m^2}$, pink about 0.01 $\frac{W}{m^2}$, dark blue about 0.1 $\frac{W}{m^2}$, light blue about 1 $\frac{W}{m^2}$, cyan about 10 $\frac{W}{m^2}$, green about 100 $\frac{W}{m^2}$, same color and marker as Fig. 3.10), and the data points correspond to the average peak value V_{out_i} reached by the 12 pixels in response to a step change in the irradiance I_{LED_i}, as shown in Fig. 3.5. The steady-state response (black points) was obtained with $I_{LED_i} = I_{LED_0}$ at several values of I_{LED_0}, whereas the initial values of the irradiance I_{LED_0} are indicated by large black circles. The shaded areas were obtained by plotting the minimum to maximum values of the mean pixel output voltages. The average dispersion of each curve (σ_{mean}) ranged from 37.2 mV in the case of the green one, to 85.4 mV, in that of the red one. (b) Average peak response of the 12 pixels versus the contrast. The various curves correspond to the S-shaped curves in Fig. 3.11(a) (same colors and markers). The contrast is given by the Michelson formula: $c_i = \frac{I_{LED_i} - I_{LED_0}}{I_{LED_i} + I_{LED_0}}$.

obtained a LSB voltage of $V_{LSB} = \frac{2.4}{2^{10}} = 2.35\,mV$, which corresponds to half of the voltage we would have obtained by setting the high reference of the ADC at V_{dd}. As the pixel output signal was limited to $[1, 2.4]\,V$, we obtained a effective resolution of the analog-to-digital conversion of $\frac{1.4\,V}{2.35\,mV} = 596 = 2^{9.22}$, i.e. about 9 bits.

The fact that the left part of the pink and red S-shaped curves is lower than that of the other ones was due to the dark current of the photodiodes. The photodiode current can be divided into 2 components: the background current I_{lbg}, which depends on the irradiance, and the dark current I_{dark}, which has a low constant value that does not depend on the irradiance. Therefore, as long as $I_{lbg} \gg I_{dark}$, we can assume that $I_{ph} \approx I_{lgb}$ and equation (3.4) still holds if we substitute I_{LED} into I_{ph}, whereas when $I_{lbg} \ll I_{dark}$, we have $I_{ph} \approx I_{dark}$, and V_{out} takes a constant value regardless of I_{LED}.

In addition, the dispersion of the S-shaped curves from one pixel to another can be seen in the shaded areas in Fig. 3.11(a). It is worth noting that this dispersion seems to be higher in the case of positive contrasts. This pattern is mainly due to the generation of the current I_b and not to the pixels themselves. As I_b is entirely conveyed to the normalizer's output current I_{out} in the case of positive contrasts, its dispersion is directly transmitted as well. This behavior can be improved by improving the dispersion of the I_b generation cell.

In Fig. 3.11(b), the data points in Fig. 3.11(a) have been plotted versus the contrast. It can be seen from this figure that the pixel's response decreased almost linearly with the luminous contrast regardless of the average luminosity, as predicted by equation (3.6). The non-linearity observed with highly negative contrasts was due to the non-linearity of the VLSI current-to-voltage converter. The contrast resolution K_c can be defined as the coefficient of c_i in (3.6) divided by 100, i.e. $K_c = -\frac{R_f I_b}{2\times 100} = -4.4\,\frac{mV}{\%}$. As the noise level is about $5\,mV$ (i.e. $2\times$LSB), contrasts as low as 2% can be detected. In addition, as the LSB voltage V_{LSB} is about 2.35mV, it is worth noting that a 1% contrast gave rise to a 2-bit change.

Faithfulness of the M^{2}APix characterization to the Michaelis-Menten model

The absolute value of the errors observed between the peak values V_{out_i} in Fig. 3.11(a) and the theoretical values $V^*_{out_i}$, based on the model for the circuit defined in Section 3.2.6, is presented in Fig. 3.12. As post-layout simulations of the circuit showed that the current-to-voltage converter gave a non-linear response at low current values, the theoretical values $V^*_{out_i}$ were computed by applying a look-up table of the current-to-voltage converter to equation (3.2), with $I_{ph} = I_{LED}$ instead of using equation (3.4).

It can be noted that on each S-shaped curve (each of which is presented in a different color), the difference between the pixel output signals and the model outputs ranges from 0.1% to 8% of the whole output range, showing a good match with the Michaelis-Menten function. Possibly due to the existence of a mismatch between the actual response of the current-to-voltage converter and the simulated post-layout response, the error decreased almost monotonically with the irradiance. In addition, the existence of a greater error in the case of the red curve ($I_{LED_0} = 0.001\,\frac{W}{m^2}$) was due to the presence of the photodiode dark current (I_{dark}), as explained above.

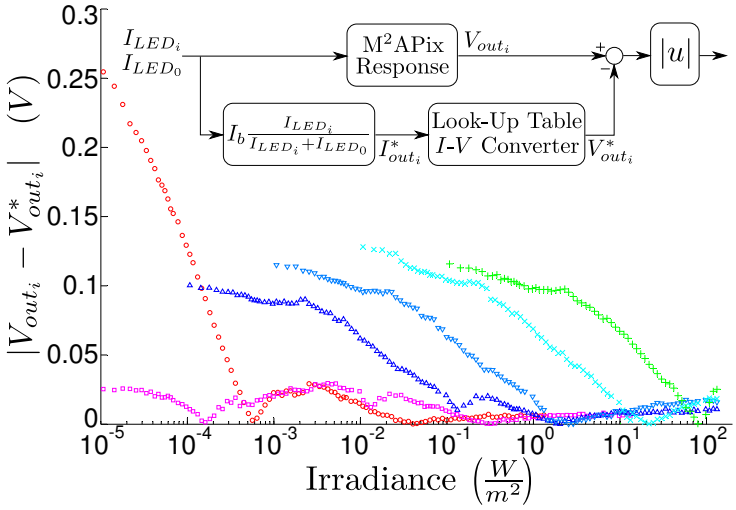

Figure 3.12: Absolute value of the error between the M^{2}APix responses measured and those predicted by the model with respect to the LED irradiance. As post-layout simulations of the circuit showed that the current-to-voltage converter gave a non-linear response at low current values, the theoretical values $V^*_{out_i}$ were computed by applying a look-up table of the current-to-voltage converter to (3.2), with $I_{ph} = I_{LED}$ instead of using equation (3.4). (Same color and marker as previous figures)

3.2.10 Comparison between M^{2}APix and Delbrück pixels

In this section, it is proposed to present a quantitative comparison between the M^{2}APix and the Delbrück pixel [175] implemented in the same silicon retina (see Fig. 3.2) when tested under the same conditions.

Table 3.1 lists all the values discussed in Section 3.2.9 for the M^{2}APix and the corresponding values obtained for the Delbrück pixel.

The Delbrück pixel reproduced non-monotonic responses with respect to the LED intensity for medium-high luminosity, showing unreliable responses for changes higher than ± 1 decade at $I_{LED_0} > 1\,\frac{W}{m^2}$. Furthermore, the steady-state response increased with the luminosity, resulting in a higher DC output variation and a lower effective output range. Finally, the adaptation time constant of the M^{2}APix, which determines the bandwidth of the output signal, can be potentially modified externally.

To show how the two types of pixel respond to small contrasts while they are still adapting to the average luminosity, we applied the procedure presented in Section 3.2.8, but contrary to the static case (see Section 3.2.9), they were exposed to repeated sequences of $0.5\,s$-long changes in the LED irradiance (corresponding to contrasts ranging from -10 to $10\,\%$) while the average irradiance was increased by 1 or 2 decades every $5\,s$.

Figure 3.13 shows the time responses of one M^{2}APix and one Delbrück pixel when exposed to two different step sequences. Both types of pixel responded quickly to small changes (small contrasts) while adapting to the average luminosity, regardless of the average luminosity and the changes in the luminosity (amounting to 1 or 2 decades in this example). However, the two pixels' responses presented some remarkable differences:

	M^{2}APix	Delbrück Pixel
Adaptation Range	7 decades	7 decades
Sensitivity Range	2 decades	1 decade
Sensitivity ($\pm$0.5 decade)	600 $\frac{mV}{Log(I)}$	600 $\frac{mV}{Log(I)}$
Absolute Contrast Resolution ($\pm$0.5 decade)	4.4 $\frac{mV}{\%}$	5 $\frac{mV}{\%}$
Effective Output Range ($\pm$1.5 decades)	1.4 V	0.8 V
DC Output Variation (7 decades)	0.1 V	0.6 V
Average Rise Time ($I_{LED_0} > 10^{-1}\,\frac{W}{m^2}$)	$\sim 1\,ms$	$\sim 1\,ms$
Maximum Rise Time ($I_{LED_0} = 10^{-3}\,\frac{W}{m^2}$)	$\sim 100\,ms$	$\sim 30\,ms$
Monotonic Response ($\pm$3 decades)	Yes	No
Adjustment of the Adaptation Time Constant	Yes	No

Table 3.1: Specifications of the M^{2}APix and Delbrück pixels. The main advantage is that M^{2}APix responds monotonically over very wide illuminance range without any loss of sensitivity and contrast resolution.

- The M^{2}APix (Fig. 3.13(b)) consistently responded to any change in the light up to ± 2 decades, while the Delbrück pixel (Fig. 3.13(c)) responded asymmetrically for positive and negative changes (see the blue line in the dotted circled area at $10\,s$) and nearly did not respond to a -2-decade change (see the red line in the dotted circled area at $15\,s$).

- The M^{2}APix always reached the same steady-state value V_{out_0} independently to the average luminosity, while the Delbrück pixel reached different steady-state values for different average luminosity.

- The contrast response produced was not the same under both light-adapted and light-adapting conditions: it depended on the change in the light. For both types of pixel, even very small contrasts (e.g. a 2% contrast) were accurately detected when the pixel had adapted to the average luminosity (see the blue line in the left-hand zoomed part of Figs. 3.13(b) and 3.13(c)). Conversely, when adapting to a sudden large change in the light, the M^{2}APix failed to detect the same 2% contrast (see the blue line in the right-hand zoomed part of Fig. 3.13(b)) due to the logarithmic compression imposed by the Michaelis-Menten function and the high fall time (T_f). During this fall time (T_f), the contrast sensitivity of the pixel increases as the slope around the operating point increases when moving from V_{out_i} to V_{out_0} (see the blue circles in Fig. 3.5). In the same condition, the Delbrück pixel shows a very different response: the first of the two consecutive 2% contrasts is detected with a much higher gain than in the light-adapted condition while the second one is nearly not detected (see the blue line in the right-hand zoomed part of Fig. 3.13(c)).

It is worth noting that the M^{2}APix responded appropriately to a 6% contrast just after a 1-decade change (see the red line in the right-hand zoomed part of Fig. 3.13(b)), which still corresponds to a good

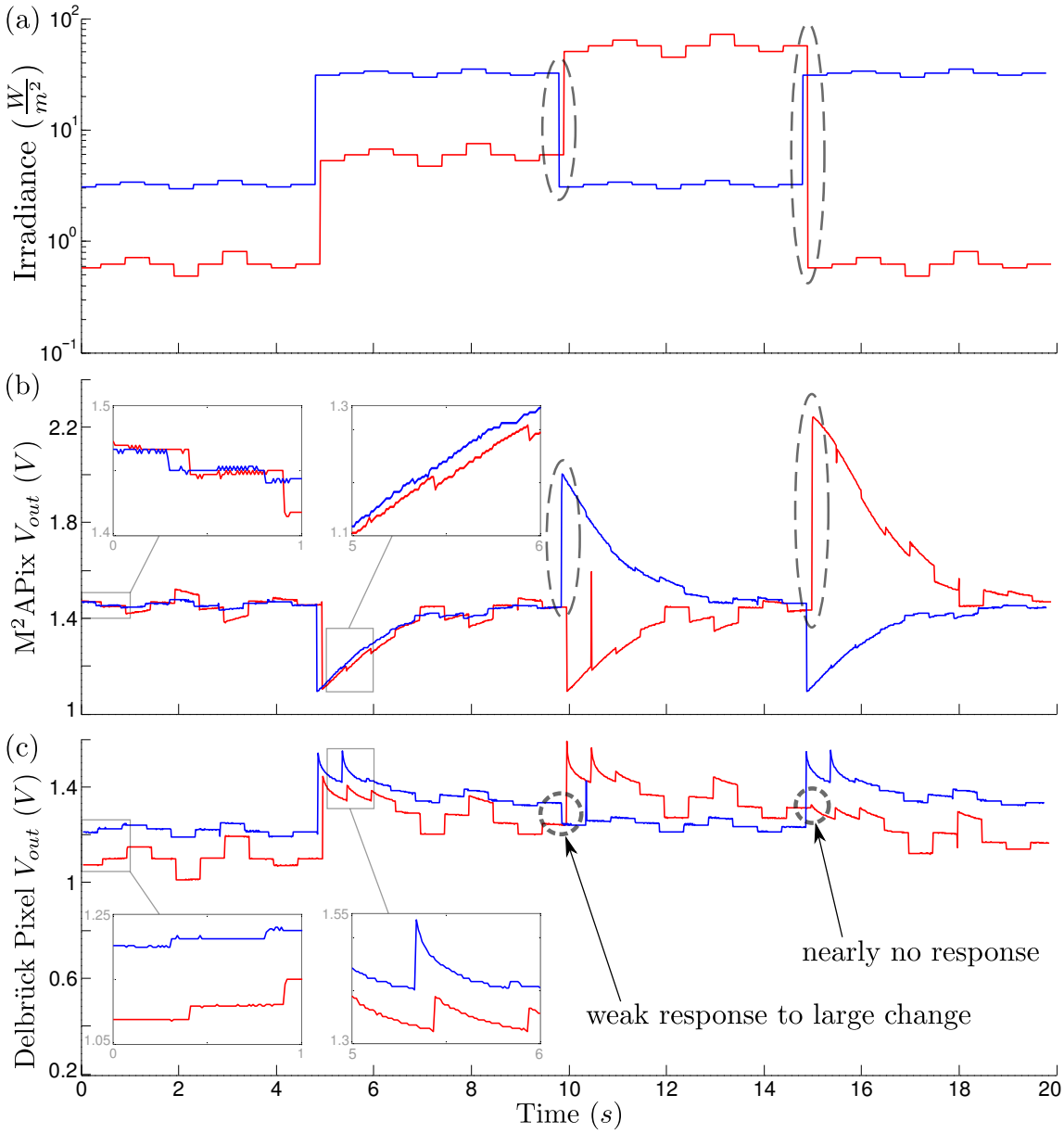

Figure 3.13: (a) Two sequences of small contrasts in a 2-decade irradiance range. Time responses of (b) M²APix and (c) Delbrück pixel [175], implemented in the same silicon retina (see Fig. 3.2), when exposed to the irradiance sequences in Fig. 3.13(a). The sequences were obtained by repeating: ±2% and ±4% contrasts (blue line), ±6% and ±12% contrasts (red line). The steps in the sequence were triggered every 0.5 s and the changes in the average irradiance every 5 s. For the sake of clarity, the timing of the sequence has been slightly shifted. (The spurious peaks, such as that which occurred at 10.5 s, may have been due to some error in the data transmission)

contrast sensitivity under light-adapting conditions. Moreover, such contrast sensitivity can be improved by decreasing the time constant of the current averaging block by reducing the value of the external capacitor C_m (see Section 3.2.6 for details). In this way, the fall time T_f would be lower and the pixel's

sensitivity would increase faster (see Section 3.2.7 for details).

3.2.11 Conclusions

The results presented here show that our Michaelis-Menten analog auto-adaptive pixel (called M^{2}APix) can adapt automatically to any irradiance from 10^{-5} to 10^2 $\frac{W}{m^2}$, which would correspond, if the LED light was green, to an illuminance of about 7×10^{-3} and 7×10^4 Lux respectively, that is from half moon on a clear night to nearly direct sunlight. At the same time, the results obtained showed that the circuit does not deviate from the model, even when the light suddenly changes by up to ± 3 decades, which would correspond, for instance, to shifting from a very dark overcast sky to direct sunlight.

In short, our auto-adaptive pixel can be said to have the following noteworthy features:

- adaptation to light in a 7-decade range, while remaining sensitive to changes in the light of up to about 2 decades;
- quasi-constant steady-state response in a 7-decade range: it produces the same contrast response whatever the average luminosity;
- no circuit deviations from the model within a ± 3-decade range of the operating current;
- constant limited-range responses at any average luminosity, resulting in a lower LSB voltage and therefore in a higher contrast resolution;
- minimum detectable contrast of 2% in the light-adapted condition and 6% in the light-adapting condition.

Some further improvements could be made in a future version of the M^{2}APix in the conversion stage, by subtracting $1\,V$ from the pixel output signal before the conversion or increasing the resistance R_f of the current-to-voltage converter in order to eventually increase the conversion range to $[0, 2.4]\,V$. This would increase the contrast resolution nearly two-fold, giving a 4-bit change in response to a 1% contrast, corresponding to a minimum detectable contrast of 1% in the light-adapted condition and 3% in the light-adapting condition.

Since the M^{2}APix makes a satisfying compromise between a high sensitivity and a wide dynamic range, it should provide a useful tool for motion detection and optical flow processing in a very large range of lighting conditions, from half-moonlight to full daylight. Thus, our auto-adaptive silicon retina, or retinas composed of larger arrays of M^{2}APix, could be employed in several fields, from event-based applications to bio-robotics and bio-medical applications.

In the future, we plan to test our silicon retina outdoors with a suitable optical lens both at night and in the daytime, by using it to measure the optic flow, for instance. In particular, we are willing to mount one or more M^{2}APix-based sensors onboard mobile and aerial robots (as in [216]) as aids to navigation under various luminosity conditions, as well as on vehicles in the case of automotive applications.

3.3 Conclusion

In the paper presented here, we showed that the novel auto-adaptive pixel, called M^{2}APix, makes a satisfying compromise between a high sensitivity (minimum detectable contrast of 2%) and a wide dynamic range (7 decades) without deviating in presence of very large changes in light (± 3 decades). The silicon retina based on the M^{2}APix pixels might therefore be used for optical flow processing in high-dynamic-range lighting conditions, from half-moonlight to full daylight.

The next step was therefore to develop a new optic flow sensor based on this auto-adaptive retina and test it in real operating conditions, for instance, in road environments. Such a new optic flow sensor, together with the experimental results obtained when using it on a car-like robot, will be presented in the next chapter.

Chapter 4

A New Bio-Inspired Optic Flow Sensor and its Application to Visual Odometry

4.1 Introduction

In chapter 2, we presented some preliminary tests performed with two LMSs based on those presented in [177,216], using a new time-of-travel method for computing the optic flow in a more robust way. The results presented showed that such LMSs could deliver relatively high-frequency measurements robustly with respect to the visual patterns and the vibrations present on the vehicle. However, some improvements were needed to obtain robust optic flow sensors that could be embedded on vehicles for use in ADAS systems: auto-adaptation to light, constant resolution of the optic flow measurements and onboard implementation of the algorithm.

In chapter 3, we presented a new bio-inspired silicon retina based on a novel pixel, called M^2APix, that reproduce the Michaelis-Menten law inspired by the findings on the animal and human retina, providing auto-adaptation to light levels in a 7-decade range and a minimum detectable contrast of about 2%.

We therefore developed a new optic flow sensor based on this auto-adaptive silicon retina, in which an improved version of the algorithm used in the work presented in chapter 2 was implemented onboard a micro-processor. The algorithm was improved in order to (i) obtain a constant optic flow resolution and (ii) reduce the computational cost, allowing a high-frequency processing while delivering measurements in a wide range and with relatively high resolution. This time, the sensors were tested on a low-cost car-like robot instead of a real vehicle mainly due to the little availability of the latter, but also to allow faster prototyping and sensors integration as well as more comfortable manipulation of the hardware used during the experimental tests.

As mentioned in sections 1.3, 1.4 and 1.5, besides the effective perception of the environment around them, autonomous vehicles also need to effectively perceived their motion to correctly estimate their trajectory (odometry) and performed closed-loop maneuvers. The sensing technologies traditionally used

for this task on mobile robots and automobiles are wheel sensors, inertial measurement units (IMUs) and global navigation satellite systems (GNSSs), such as standard or differential GPSs, but these sensors often suffer from drift, low resolution, high noise levels or limited applicability to specific environments [278,279]. In fact, apart for the GNSSs that give directly the vehicle's absolute position but only outdoors and with low resolution, none of these sensors can directly provide either position or velocity information with respect to a local inertial frame. Several approaches based on visual odometry have therefore been recently developed using various visual sensors, such as standard cameras, optical-mouse sensors and custom-made optic flow sensors (see section 1.9). A remarkable example of the potential interest of visual sensors for measuring the vehicle's longitudinal and lateral velocity is given by the Correvit optical sensors [280], which are used, for instance, on race cars as well as on prototype vehicles as ground-truth reference values [281].

In the following paper, we will present (i) the new optic flow sensor based on the M^2APix pixel and the new optic flow algorithm, and (ii) its application to ground visual odometry on a low-cost car-like robot, called BioCarBot, which stands for Bio-inspired visually-guided Car-like Robot. The experimental results obtained indoors and outdoors in several operating conditions (light levels, ground textures, vibrations, etc) are shown and discussed in details, and they are also compared with those obtained using measurements based on an Inertial Measurement Unit (IMU) and a motor's speed sensor.

The details on the new algorithm implementation and on the electronics used on the BioCarBot robot, as well as the supplementary data of the article, are given in Appendix C.

4.2 Article

Title: "Minimalistic optic flow sensors applied to indoor and outdoor visual odometry on a car-like robot"

Article submitted in February 2016 in *Bioinspiration & Biomimetics*

Authors: Stefano Mafrica,[1, 2] Alain Servel[2] and Franck Ruffier[1]

4.2.1 Abstract

Here we present a novel bio-inspired optic flow (OF) sensor and its application to ground visual odometry on a low-cost car-like robot called BioCarBot. The minimalistic OF sensor was robust to high-dynamic-range lighting conditions and to various visual patterns encountered thanks to its M^{2}APIX auto-adaptive pixels and the new cross-correlation OF algorithm implemented. The low-cost car-like robot estimated its velocity and steering angle, and therefore its position and orientation, via an Extended Kalman Filter (EKF) using only two downward-facing OF sensors and the Ackerman steering model. Indoor and outdoor experiments were carried out, in which the robot was driven in the closed-loop mode based on the velocity and steering angle estimates. The experimental results obtained show that our novel OF sensor can deliver high-frequency measurements ($> 300\,Hz$) in a wide OF range (1.5 to $15\,\frac{rad}{s}$) and in a 7-decade high-dynamic light level range. The OF resolution was constant and could be adjusted as required (up to $0.05\,\frac{rad}{s}$), and the OF precision obtained was relatively high (standard deviation of $0.17\,\frac{rad}{s}$ with an average OF of $4.5\,\frac{rad}{s}$, under the most demanding lighting conditions). An EKF-based algorithm gave the robot's position and orientation with a relatively high accuracy (maximum errors outdoors at a very low light level: $0.95\,m$ and $0.58\,rad$ over about $32\,m$ and $8\pi\,rad$) despite the low-resolution control systems of the steering servo and the DC motor, as well as a simplified model identification and calibration. The minimalistic OF-based odometry results were lastly compared to those obtained using measurements based on an Inertial Measurement Unit (IMU) and a motor's speed sensor.

4.2.2 Introduction

Optic flow (OF) has been studied by many authors during the last few decades [100, 109, 282, 283]: various OF sensors [177, 192, 209] and algorithms [98, 99, 111] have been developed and used for robotic applications such as autonomous navigation [233, 284], speed control [242, 285], simultaneous localization and mapping (SLAM) [21, 22], and visual odometry (VO) [286–289].

Although standard cameras have been widely used in this context, they have often proved to be unsuitable for OF computation purposes, especially outdoors, because of their low dynamic range, their

[1] Aix-Marseille Université, CNRS, ISM UMR 7287, 13288 Marseille, France
[2] PSA Peugeot Citroën, Route de Gisy, 78140 Vélizy-Villacoublay, France

low frame rate and the high computational cost of the image processing. In addition, in camera-based VO and SLAM applications, small motion is usually assumed to occur between frames, which is often not the case, and a "scale factor" is also required to convert the visual information (e.g. the OF) into metric measurements, which is often provided by other sensors [289, 290]. These issues can be overcome by using high-dynamic-range, high-frame-rate cameras and stereo-vision systems, but these solutions require even larger computational resources and are too expensive for most robotic and automotive applications.

Ethological findings on insects have shown that complex navigation tasks such as terrain following [291, 292] ([239] for a review), speed control [92] and odometry [146] are performed by many flying insects, such as flies and bees, on the basis of OF cues although their compound eyes have a very poor spatial resolution in comparison with our modern high-resolution cameras. In addition, it has been behaviorally shown that nocturnal insects use these OF cues also for flight control in dim light conditions [293].

Custom-made bio-inspired visual sensors have therefore been developed using technologies of various kinds [183, 219, 294, 295] in order to deal with the issues encountered with standard cameras. However, designing effective low-cost OF sensors is a real challenge, and additional information given by other sensors or based on assumptions about the environment is still required to convert the OF to distance and/or velocity values. In addition, OF alone is generally thought to be unsuitable for VO purposes when working in complex environments where the sweeping angular speed between the sensor and the objects that induce the OF is subjected to sharp, large variations, especially in the presence of occluded objects [289, 290].

Visual odometry for mobile robots and automobiles has been recently developed using downward-facing standard cameras [296–300] and optical mice [301–303], as well as custom-made OF sensors [304, 305], since the visual patterns and light conditions encountered in this way are relatively uniform, and the distance between the sensors and the ground usually changes slightly. In most cases, the authors used the car-like non-holonomic constraint to estimate the vehicle's position and orientation using simple methods involving low computational costs.

However, solutions based on standard cameras [297, 300] still fail to cope with high-dynamic-range lighting conditions, as well as being impeded by the low frame rate and the high computational cost of the image processing: only a narrow range of low velocity measurements can often be obtained using this approach.

Solutions based on optical mice [301, 303] are certainly very cheap and deliver high-frequency measurements, but their main disadvantage is that they have to operate very near the ground to be able to work properly, and are therefore unsuitable for use in environments with an uneven terrain. In addition, these sensors are usually highly sensitive to the lighting conditions and like standard cameras, deliver measurements in a rather narrow velocity range.

Solutions based on custom laser or LED-lighted OF sensors [304, 305], have been developed in order to reduce the sensitivity to height and improve the performances while traveling over terrains of various

kinds, but no tests were run by these authors under various lighting conditions and under the robot's normal driving conditions.

Physiological studies on vertebrates have shown that the retina obeys a process of adaptation whereby each photoreceptor's response is normalized by a representative value of the average local luminosity, in line with the Michaelis-Menten equation [36, 47]. In the light of these and subsequent findings, many efforts have been made to mimic the Outer Plexiform Layer (OPL) circuitry in silicon retinas [171, 175], or to implement the model in software for image processing [179, 272]. More recently, we presented a novel auto-adaptive pixel called M2APix, which stands for Michaelis-Menten Auto-Adaptive Pixel, that auto-adapts in a 7-decade range and responds appropriately to both small and large contrasts [261].

Several versions of Local Motion Sensors (LMSs) have been developed at our laboratory [122, 177, 216, 242]: local 1-D OF measurements have been obtained using very few pixels and applying methods based on findings previously obtained on the fly's visual system [62].

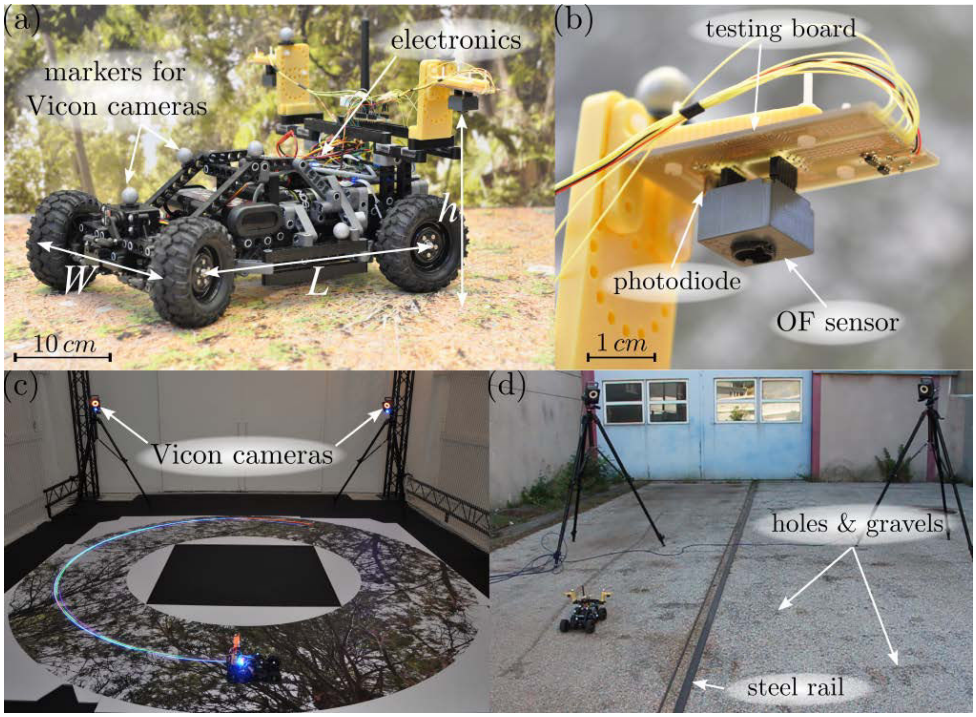

Figure 4.1: (a) Picture of the low-cost car-like robot called BioCarBot. (b) Picture of one of the 2 OF sensors used on the robot. (c) Picture of the indoor test environment equipped with Vicon cameras. (d) Picture of the outdoor test environment equipped with Vicon cameras. The ground, which consisted mainly of asphalt, included holes, gravel and a steel rail.

In this paper, we present:

- a novel bio-inspired OF sensor giving measurements which are robust to high-dynamic-range lighting conditions and to the various visual patterns encountered;
- an application to visual odometry, in which these minimalistic OF sensors were mounted on a

low-cost car-like robot called BioCarBot (see Fig. 4.1) and tested both indoors and outdoors in a 7-decade light-level range.

An Extended Kalman Filter (EKF) was used to estimate the robot's velocity and steering angle using only the OF measurements delivered by two downward-facing sensors, as presented in [306] in the case of a single indoor environment. In the present paper, the tests were extended to various indoor (see Supplemental Video 1) and outdoor (see Supplemental Video 2) testing conditions, including various light levels, ground textures, trajectories and vibrations, and the robot's absolute position and orientation were estimated in real time using the EKF estimates and the Ackerman steering model. These position and orientation estimates were lastly compared with those obtained by applying the same model to the measurements from an Inertial Measurement Unit (IMU) and a motor's speed sensor.

The results presented here show that our novel OF sensors were robust to changes in light in a 7-decade range (from about $10^{-10}\,A$ to $10^{-3}\,A$ of the photodiode current), including sharp changes of up to 2 decades occurring within $0.5\,s$. Although low-cost, low-resolution servos and motors and a simplified model identification and calibration method were used, the robot was able to estimate its velocity and steering angle accurately. The robot's position and orientation were estimated both indoors and outdoors while it was traveling through unstructured environments (on ground consisting of asphalt, gravel, sand, and leaves, including shadows and holes), both during the day and at night. In addition, the visual odometry method was robust to vibrations liable to change the sensors' local height by up to 6% (i.e. $\pm 10\,mm$ over $175\,mm$).

In Sec. 4.2.3, we will introduce the principles underlying a 2-pixel LMS as well as the method and the hardware used to construct our novel OF sensor. In Sec. 4.2.4, we will present the BioCarBot robot, the simplified model used to estimate the robot's velocity and steering angle, and the control scheme implemented on the robot. In Sec. 4.2.5, we will present and discuss the results obtained in the indoor and outdoor experiments performed. Some conclusions will be reached in the last section.

4.2.3 New Implementation of the Visual Motion Sensor

Principles Underlying a 2-Pixel Local Motion Sensor

A defocused lens placed in front of 2 photoreceptors determines the interreceptor angle $\Delta\varphi$ between the 2 photoreceptors' axes and gives them a Gaussian angular sensitivity with an acceptance angle $\Delta\rho$ (Fig. 4.2(a)), on similar lines to what occurs in many insects' eyes. A visual contrast moving in front of the LMS will induce a time lag τ between the photoreceptors' output signals (Fig. 4.2(b)). After measuring this time lag, the optic flow can be computed as follows:

$$\omega(t) = \pm\frac{\Delta\varphi}{\tau(t)}, \tag{4.1}$$

where the sign depends on the orientation of the sensor's axis and on which of the two output signals is delayed.

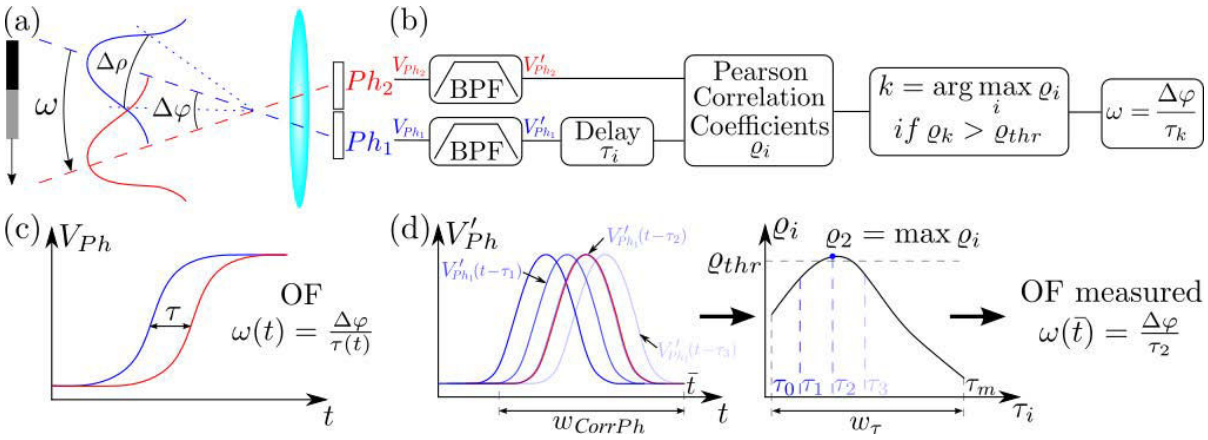

Figure 4.2: (a),(c) Principles underlying a 2-pixel LMS. The 1-D optic flow ω produced by a contrast moving in front of the LMS can be computed taking $\omega(t) = \frac{\Delta\varphi}{\tau_k(t)}$, where $\Delta\varphi$ is the interreceptor angle and τ is the time lag between the two photoreceptors' output signals. (b),(d) Block diagram of the novel OF algorithm that determines the time lag τ_k between several τ_i in a time window w_τ, giving the maximum cross-correlation between the delayed and non-delayed signals. The OF sensor's output does not result directly from the correlation, as in the Reichardt-Hassenstein model [76], but from the ratio between the interreceptor angle $\Delta\varphi$ and the time lag τ_k, i.e. $\omega(t) = \frac{\Delta\varphi}{\tau_k(t)}$.

The acceptance angle $\Delta\rho$, namely the full width at half-height of the Gaussian angular sensitivity, determines the cut-off frequency of the low-pass spatial filter as follows: $F_c \approx \frac{1}{\Delta\rho}$ [307]. On the one hand, achieving a tight $\Delta\rho$ makes it possible for the photoreceptors to respond to higher spatial-frequency contrasts but on the other hand, when $\Delta\rho < \Delta\varphi$, it is difficult to estimate the time lag τ because the 2 output signals are rarely correlated in a short time window. Therefore, as occurs in some diurnal insects [61], we adjusted the distance between the plane of the lens and that of the photoreceptors in order to obtain $\Delta\rho = \Delta\varphi$.

New Cross-Correlation Method for Computing the OF

To compute the 1-D OF ω, the time lag τ between 2 neighboring pixels' output signals was estimated using a cross-correlation method inspired by the Reichardt-Hassenstein correlator model [76]. In the method presented here, the OF sensor's output did not result directly from the correlation, as in the Reichardt-Hassenstein model, but from the ratio between the interreceptor angle $\Delta\varphi$ and the time lag τ_k giving the maximum cross-correlation between the delayed and non-delayed signals, i.e. $\omega(t) = \frac{\Delta\varphi}{\tau_k(t)}$, on similar lines to the method proposed in [123].

First the 2 pixels' output signals (V_{Ph_1}, V_{Ph_2}) were sampled and band-pass filtered (V'_{Ph_1}, V'_{Ph_2}) at $f_l = 3\,Hz$ and $f_h = 30\,Hz$, and the following pseudo-algorithm was then applied (Fig. 4.2(c-d)):

1. delay one of the two filtered signals (e.g. $V'_{Ph_1}(t)$) by the time τ_i;

2. compute the Pearson correlation coefficients between the delayed (e.g. $V'_{Ph_1}(t-\tau_i)$) and non-

delayed signals (e.g. $V'_{Ph_2}(t)$) in a fixed time window w_{CorrPh};

3. repeat step 1 and 2 for every τ_i ($i = 1, \ldots, n$) in a fixed time window w_τ;

4. set τ at a value equal to the time lag τ_k giving the maximum cross-correlation coefficient, as long as this maximum is greater than the fixed value ρ_{thr};

5. compute the OF ω using the equation (4.1) (or set ω at NaN if the maximum cross-correlation coefficient is less than ρ_{thr}).

The threshold value of the cross-correlation coefficients (ϱ_{thr}) directly reflects the reliability and the robustness of the OF measurements: the higher ϱ_{thr} is, the more reliable and robust the measurements will be, but in the presence of noise, the lower the refresh rate will be.

The time window of the signals (w_{CorrPh}) determines the bandwidth of the OF measurements, whereas the number of samples (n) in the time window determines the reliability of the correlation coefficients: the smaller w_{CorrPh} is, the larger the bandwidth will be, and the higher n is, the higher the reliability will be.

In order to obtain a constant resolution $\Delta\omega$, the time lags were chosen as follows:

$$\tau_i = \frac{\Delta\varphi}{|\omega_i^*|}, \tag{4.2}$$

where ω_i^* are the desired OF measurements, which are linearly separated by the resolution $\Delta\omega^*$ required. The signals can then be delayed by the time τ_i elapsing between two sampling steps, by linearly interpolating the signals sampled.

Hardware and Software Implementation

In this study, we used the auto-adaptive silicon retina presented in [261] soldered onto a tiny printed circuit board (PCB) on which an optical lens casing was mounted (Fig. 4.3).

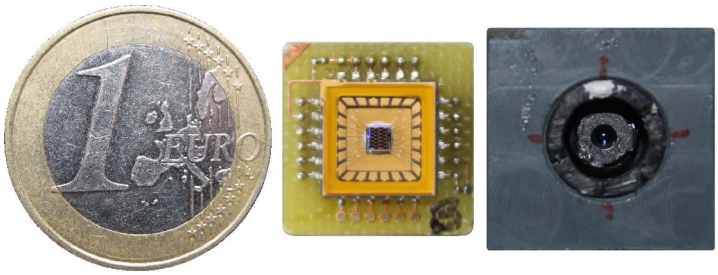

Figure 4.3: Picture of the auto-adaptive silicon retina comprising the 12 M^{2}APix pixels [261] soldered onto a tiny PCB on which the optical lens casing was mounted.

The M^{2}APix pixel, which stands for Michaelis-Menten Auto-Adaptive Pixel, can auto-adapt in a 7-decade range and responds appropriately to small contrasts, such as $\pm 2\%$, as well as large changes in light, such as ± 3 decades [261]. In the chip used here, the analog low-pass filter had a cut-off frequency

of $100\,Hz$ (instead of $300\,Hz$ used in [261]), giving a minimum sampling frequency of $200\,Hz$ in order to prevent the occurrence of aliasing.

The optical lens used here was taken from a Raspberry-Pi camera (focal length: $2\,mm$), while the lens casing was custom made using a 3D printer to precisely adjust the distance between the plane of the lens and that of the pixels during the calibration phase. The interreceptor angle $\Delta\varphi$ and the acceptance angle $\Delta\rho$ were measured at $\Delta\varphi \approx \Delta\rho \approx 3.6^o$ using the method presented in [177], giving a cut-off frequency of the low-pass spatial filter $F_c \approx \frac{1}{\Delta\rho} \approx 0.28 \frac{cycles}{deg}$ [307].

The OF algorithm presented in section 4.2.3 was implemented by setting:

- the threshold on the cross-correlation coefficients $\varrho_{thr} = 0.99$;
- the number of pixels signals samples $n = 70$, giving a signal time window for the cross-correlation computation $w_{CorrPh} = 0.21\,s$;
- the number of time lags $m = 30$, because a larger number would cause saturation of the CPU load since the implementation of the algorithm was not optimized.

Depending on the velocity range required, the time lag window w_τ ranged from $6.3\,ms$ to $63\,ms$ in order to obtain OF measurements ω_i ranging from 1 to $15\,\frac{rad}{s}$ and a resolution $\Delta\omega^*$ ranging from 0.05 to $0.5\,\frac{rad}{s}$. In the experiments presented here, the OF range was set prior to each test at the smallest range comprising all the reachable OF values given the robot's velocity and steering angle commands and the sensors' height, in order to obtain the highest OF resolution given the computational constraint on the total number of possible time lags at each time step ($m = 30$).

We note that having a very high threshold on the cross-correlation coefficients ($\varrho_{thr} = 0.99$) does not imply that the 2 neighboring signals have to be identical but rather that they have to be linearly dependent, i.e. there might be a gain and an offset between them. Thus, the output of our OF algorithm, i.e. the OF measurements, nearly does not depend on the color, intensity and spatial frequency of visual patterns encountered, i.e. on the amplitude and shape of the pixels' temporal signals produced, as long as these signals fall in the bandwidth of the band-pass temporal filter (see section 4.2.3). In other words, we can say that the pattern-based noise is very low as long as the visual patterns are not completely uniform and have spatial frequency components lower than F_c.

The OF algorithm was then applied to every pair of adjacent pixels in each of the two 6-pixel rows (see Fig. 2(b) in [261]), giving 10 local 1-D OF measurements within a field of view of about 18^o. The Visual Motion Sensor (VMS) was therefore composed of 10 2-pixel LMSs and the median value ω_m of the 10 OF measurements ω_i was used as the actual output of the VMS to robustly filter out possible outliers.

4.2.4 BioCarBot: a Bio-Inspired Visually-Guided Car-Like Robot

The Car-Like Robot

Figure 4.1 shows a picture of: the BioCarBot robot, one of the 2 VMSs used and the indoor and outdoor testing environments equipped with a Vicon motion-capture system.

The present odometry method was tested using a car-like robot based on the 2WD Racecar Kit provided by *Minds-I Robotics*, which was chosen despite the low resolution of the servo and motor control because of its modularity and low price (US$ 275). The car-like robot was composed of a 1/10-scale car body ($419 \times 203 \times 114\,mm$), one Hitec HS-311 standard servo coupled to a steering hub, one 5000-rpm DC motor connected to a 300-A electronic speed controller (ESC), one 7.2-Volt 3000-mAh Ni-Cd rechargeable battery, one mechanical slip differential and four $90mm$-diameter crawler wheels.

The embedded electronics included one Nanowii board (*Flyduino*) featuring a ATmega32u4 16-MHz CPU microcontroller (*Atmel*) and a MPU-6050 inertial measurement unit (IMU) comprising a 3-axis gyroscope and a 3-axis accelerometer (*InvenSense*) and one Overo IronSTORM computer-on-module (COM) (*Gumstix*) featuring a 1-GHz CPU DM3730 processor (*Texas Instruments*) comprising an ARM Cortex-A8 architecture and a C64x digital signal processor (DSP).

Thanks to the modularity of the robot's structure, 2 identical VMSs (Fig. 4.3(b)) were attached to the robot's frame on both sides of its body, aligned with the rear wheel axis (Fig. 4.1(a)). To facilitate the sensors' installation, we used the same testing board as that which was used in the study presented in [261] to connect the VMSs to the Nanowii board (Fig. 4.1(b)). An OSRAM BPX65 photodiode connected to an analog amplifier circuit was also included on the testing board next to the VMS in order to measure the effective light levels of the scene.

To obtain the ground-truth values, the 3-D robot's pose ($[X\,Y\,Z\,\alpha\,\beta\,\gamma]^T$) was measured by means of a Vicon motion-capture system thanks to the infrared markers attached to the robot's frame (Fig. 4.1(a)). Indoor experiments were performed in the flying arena at our laboratory [308], whereas 4 individual Vicon cameras each mounted on a tripod were used outdoors (Fig. 4.1(c-d)).

Car-Like Robot Modeling

As the robot's velocity was relatively low and the robot did not have any suspension system, we focused here on the 2-D kinematic model for a car-like robot moving on a flat surface. Figure 4.4 shows the kinematic diagram of the BioCarBot with the 2 VMSs installed on both sides, as depicted in Fig. 4.1(a).

Let us take the inertial frame $< I >$ having the x and y axes lying on the local ground plane, with the robot's body frame $< B >$ placed in the middle of the rear wheels' axis. Two VMSs were placed at $\boldsymbol{x}_l = [x_l\,y_l\,z_l]^T$ and $\boldsymbol{x}_r = [x_r\,y_r\,z_r]^T$ with respect to $< B >$ ($x_l = x_r = 0\,mm$, $y_l = -y_r = 140\,mm$, $z_l = z_r = 125\,mm$), respectively, facing downwards at a height of h_l, h_r, respectively, from the ground ($h_l = h_r = h = 175\,mm$) (Fig. 4.1(a),4.4(a)). As the sensors' frames $< l >, < r >$ were taken to be parallel to the body frame $< B >$, we can consider all the position and velocity vectors projected onto $< B >$.

The ground OF can be measured using the method presented in section 4.2.3 thanks to the no-skidding assumption, which guarantees that $V_x >> V_y$ with any velocity vector $\boldsymbol{V} = [V_x\,V_y\,0]^T$ located near the line passing through the two rear wheel/ground contact points. Therefore, the OF measured between the

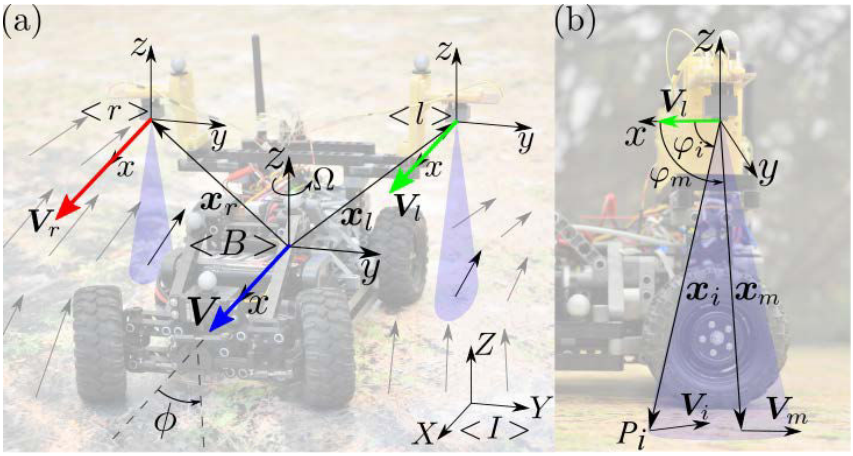

Figure 4.4: (a) Kinematic diagram of the robot moving on a flat surface. (b) Kinematic diagram of one OF sensor.

$i-1$-th and i-th pixels of each VMS can be written as follows:

$$\omega_i = -\frac{V_{i_x} \sin^2 \varphi_i}{h}, \tag{4.3}$$

where $\boldsymbol{V}_i = [V_{i_x}\, V_{i_y}\, 0]^T$ is the velocity vector of the vector $\boldsymbol{x}_i = [x_i\, y_i\, z_i]^T$ ($z_i \approx -h$) giving the position of the intersection point P_i between the i-th pixel axis and the ground plane with respect to the sensor's frame, i.e. $< l >$ or $< r >$, and φ_i is the angle between $\boldsymbol{x}_i$ and the x-axis of the sensor's frame (Fig. 4.4(b)).

Since the sensors' frames$< l >, < r >$ were taken to have their x-axis aligned with the rows of pixels of each VMS (i.e. $y_i \approx 0$) and to be parallel to $< B >$ (Fig. 4.4(a),(b)), the position vectors of the intersection points P_i with respect to $< B >$ can be written as $\boldsymbol{x}_i^l = \boldsymbol{x}_l + \boldsymbol{x}_i = [x_i\, y_l\, z_l - h]^T$ and $\boldsymbol{x}_i^r = \boldsymbol{x}_r + \boldsymbol{x}_i = [x_i\, y_r\, z_r - h]^T$, for the left and right sides, respectively, and their corresponding velocities are given as follows:

$$\begin{cases} V_{i_x}^l = V_l = V - y_l \Omega \\ V_{i_x}^r = V_r = V - y_r \Omega \end{cases}, \tag{4.4}$$

where V, Ω are the robot's longitudinal and angular velocity, respectively (Fig. 4.4(a)).

By combining equations (4.3) and (4.4), we obtain a set of redundant linear equations that relate the local OF measurements ω_i to the robot's longitudinal and angular velocity V, Ω, which can be used for odometry purposes with any wheeled robot that satisfies the no-skidding assumption, regardless how the robot is actuated. However, if we are interested in controlling the robot in closed loop, the dynamic equations relating V, Ω to the specific control parameters should be included into equation (4.4) for better tracking performances.

Since we are dealing here with a car-like robot which is controlled by a DC motor, giving the longitudinal velocity V, and a steering servo, giving the steering angle ϕ, the angular velocity Ω can be computed as follows, according to the Ackermann steering geometry [264] (Fig. 4.4(a)):

$$\Omega = \frac{\tan\phi}{L} V, \tag{4.5}$$

where $L(= 255\,mm)$ is the distance between the rear and front wheel axes (Fig. 4.1(a)).

By substituting the median value of the OF measurements obtained with each VMS ω_m^l, ω_m^r into (4.3) and combining equations (4.3), (4.4) and (4.5), the following equation relating the output $\boldsymbol{\zeta} = [\omega_m^l\ \omega_m^r]^T$ to the state $\boldsymbol{\xi} = [V\ \phi]^T$ is obtained:

$$\boldsymbol{\zeta} = \begin{bmatrix} \omega_m^l \\ \omega_m^r \end{bmatrix} \approx \begin{bmatrix} \frac{(L - y_l \tan\phi)\sin^2\varphi_m^l}{hL} \\ \frac{(L - y_r \tan\phi)\sin^2\varphi_m^r}{hL} \end{bmatrix} V = \boldsymbol{h}(\boldsymbol{\xi}), \tag{4.6}$$

where φ_m^l and φ_m^r are the orientation of the pixel's axis corresponding to the median OF values ω_m^l and ω_m^r, respectively.

Lastly, the dynamics of V and ϕ, which mostly depend on the dynamics of the DC motor and the steering servo, respectively, were identified in the form of two independent first-order systems using the ground-truth measurements:

$$\dot{\boldsymbol{\xi}} \approx A\boldsymbol{\xi} + B\boldsymbol{u} = \boldsymbol{f}(\boldsymbol{\xi}, \boldsymbol{u}). \tag{4.7}$$

where $A = diag(a_1, a_2)$, $B = diag(b_1, b_2)$. The values of A and B were identified using the slower time constants as follows: $a_1 = -b_1 = -2.15$, $a_2 = -b_2 = -4.87$.

It is worth noting that the identified model given in equation (4.7) does not take into account the non-linearities which characterize the conversion from the actual control inputs, i.e. the values delivered to the ESC and the servo controller, to the model input $\boldsymbol{u}$, such as, for instance, the dependency of the DC motor's speed to the battery charge level and the backlash in the steering system's geometry. Such a simplified model, therefore, would not be accurate enough to be used directly for odometry purposes, however we want to show here that it is sufficient to obtain good odometry results when applied to an Extended Kalman Filter (EKF) using the measurements delivered by our novel OF sensors.

EKF and Control System

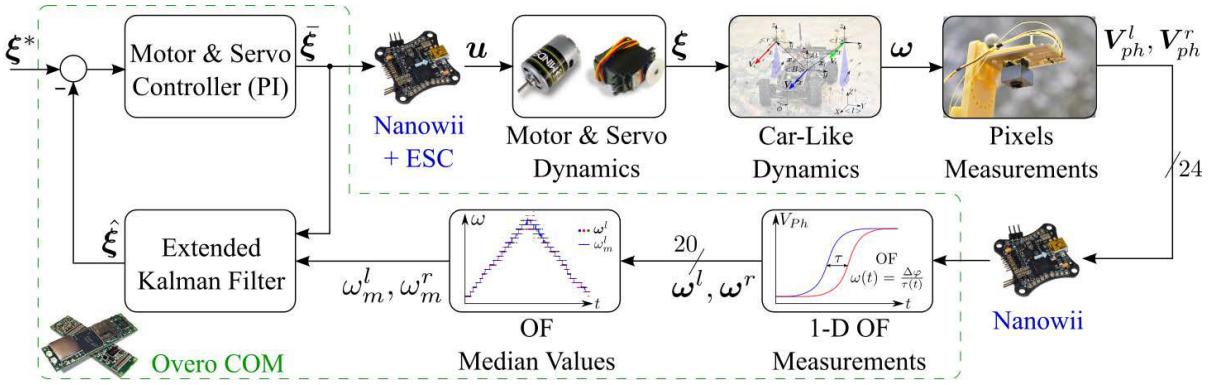

Figure 4.5: Block diagram of the estimation and control scheme implemented on the BioCarBot robot.

In order to obtain a robust continuous estimation of the robot's longitudinal velocity and steering angle (V, ϕ), an Extended Kalman Filter (EKF) based on the discrete approximation of the model presented in (4.7),(4.6) was implemented, taking the median values of the 10 local OF measurements (ω_m^l, ω_m^r) to be actual measurements.

The first-order discrete approximation for the model presented in (4.7) was taken to be as follows:

$$\begin{cases} \boldsymbol{\xi}_k = \hat{\boldsymbol{f}}(\boldsymbol{\xi}_{k-1}, \boldsymbol{u}_{k-1}, \boldsymbol{w}_{k-1}) = \\ \quad = [\boldsymbol{f}(\boldsymbol{\xi}_{k-1}, \boldsymbol{u}_{k-1}) + \boldsymbol{w}_{k-1}]\Delta t + \boldsymbol{\xi}_{k-1} \quad , \\ \boldsymbol{\zeta}_k = \hat{\boldsymbol{h}}(\boldsymbol{\xi}_k, \boldsymbol{\nu}_k) = \boldsymbol{h}(\boldsymbol{\xi}_k) + \boldsymbol{\nu}_k \end{cases} \tag{4.8}$$

where the index k denotes the k-th sampling period (i.e., $t = k\Delta t$); $\boldsymbol{w}, \boldsymbol{\nu}$ denotes the model and the measurement noise, respectively, and they are assumed to be independent white noises and to have normal probability distributions, i.e. $p(w) \sim N(0, Q)$ and $p(\nu) \sim N(0, R)$, where $Q = diag(\sigma_V^2, \sigma_\phi^2)$ and $R = diag(\sigma_l^2, \sigma_r^2)$ are covariance matrices. The assumption that Q and R had normal and uncorrelated distributions was adopted on the basis of what was observed statistically during several experimental tests with trajectories of various kinds (see section 4.2.5). The elements of R and Q were set at about $(\frac{\Delta\omega^*}{2})^2$ and 0.01^2, respectively, for the indoor tests and about $\Delta\omega^*$ and 0.02^2, respectively, for the outdoor tests.

Because of the backlash and other uncertainties, we did not have a very clear picture of the motor and steering control inputs $\boldsymbol{u}$ to the system. The real system inputs were therefore assumed to be equal to the outputs from the controller, i.e. $\boldsymbol{u}_{k-1} = \bar{\boldsymbol{\xi}}_{k-1}$ (see equation (4.9)).

The initial estimate of the state $\hat{\boldsymbol{\xi}}_0$ was set at zero, while the initial estimate of the error covariance matrix P_0 was set at the identity matrix. When there were no measurements available on at least one of the robot's sides, i.e. no ω_m^l or ω_m^r, the Kalman gain K_k was set at zero, so that it was still possible to have an estimation of $\boldsymbol{\xi}_k$ based on the "a priori" prediction $\hat{\boldsymbol{\xi}}_k^-$. In that case, a timeout was set at $0.5\,s$, after which the EKF was reinitialized and the robot was stopped. Such an event happened only in the few cases where the robot drove for a while on a non-textured area of the floor (e.g. the white or black areas in Fig. 4.1(b)), which never happened in the tests presented here.

The robot's longitudinal velocity and steering angle ($\boldsymbol{\xi} = [V,\, \phi]^T$) were controlled in the closed-loop mode using their estimates ($\hat{\boldsymbol{\xi}} = [\hat{V}, \hat{\phi}]^T$) and the values required ($\boldsymbol{\xi}^* = [V^*,\, \phi^*]^T$) via a proportional and integral (PI) controller (Fig. 4.5):

$$\bar{\boldsymbol{\xi}} = K_P(\boldsymbol{\xi}^* - \hat{\boldsymbol{\xi}}) + K_I \int (\boldsymbol{\xi}^* - \hat{\boldsymbol{\xi}}), \tag{4.9}$$

where $K_P = diag(0.5, 0.4)$ and $K_I = diag(2.5, 2)$.

We note that the EKF and the PI controller were applied on the state $\boldsymbol{\xi} = [V,\, \phi]^T$ instead of $\boldsymbol{\xi} = [V,\, \Omega]^T$ in order to obtain a more robust estimate of ϕ, and therefore better closed-loop trajectories. Also, a feedforward term could be added to the PI controller to increase the tracking performances/accuracy, although it was not included here since no significant improvements were obtained.

4.2.5 Experimental OF measurements and visual odometry results

Experiments were carried out both indoors (see Online Video 1) and outdoors (see Online Video 2) using various floor patterns and trajectories to test the performances of the OF sensors as well as the method of estimation and control presented here.

Figure 4.1 shows two examples of test environments, indoors (Fig. 4.1(c)) and outdoors (Fig. 4.1(d)) at the flying arena at our laboratory. In the tests presented here, the car-like robot followed paths forming a circle, a square and a figure of eight. The floors used in the indoor tests showed patterns of various colors and contrasts, from small and large high-frequency contrasts to large low-frequency contrasts. The ground used in the outdoor tests, which consisted mainly of asphalt, included holes, gravel and a steel rail.

The M^{2}APix output signals were sampled at a frequency of 333Hz ($\Delta t = 3\,ms$) by the on-chip ADC (see [261] for details), acquired by the Nanowii board via SPI communication and transmitted to the computer-on-module (COM) via serial communication (see section 4.2.4 for details). The OF algorithm presented in Sec. 4.2.3 as well as the estimation and control scheme shown in Fig. 4.5 were run at the same rate on the COM. The Linux-based program running on the COM was entirely generated in the Matlab/Simulink environment using the RT-MaG toolbox [308], a custom-made toolbox for real-time applications developed at our laboratory. The ground host-PC program conveys the control set points V^*, ϕ^* to the robot's COM and receives data from the robot's COM and the Vicon system using the RT-MaG toolbox. Estimates of the absolute robot's position $\hat{\boldsymbol{X}}$ and orientation $\hat{\theta}$ were lastly computed in real time by integrating the equations provided in the Ackermann model [264] using the EKF estimates $\hat{V}$, $\hat{\phi}$ as inputs to the model.

To compare the minimalistic odometry results obtained when using the OF measurements with those obtained when using inertial measurements, the estimates of the absolute robot's position and orientation $\hat{\boldsymbol{X}}_{IMU}$, $\hat{\theta}_{IMU}$ were also computed offline using the measurements acquired from the MPU-6050 IMU ($[\boldsymbol{a}_{IMU}^T\, \boldsymbol{\Omega}_{IMU}^T]^T = [a_x\, a_y\, a_z\, \Omega_x\, \Omega_y\, \Omega_z]^T$) and the measurements of the DC motor's speed Ω_{DC} acquired by a hall-effect sensor attached to the motor's shaft. Several methods have been proposed through the years to perform odometry using inertial measurements based on specific dynamic models of the IMU used, in particular to handle the varying bias of the gyroscopes [309, 310], but these methods require an accurate identification and calibration of the models used. Here, in order to have a meaningful comparison with our minimalistic OF-based method, estimates of the robot's velocity and steering angle $\hat{V}_{IMU}$, $\hat{\phi}_{IMU}$ were obtained by applying an EKF to the same dynamical model, e.g. equation (4.7), while using the following output model:[3]

[3]Note that in the first equation in (4.10) there is no contribution of $\boldsymbol{\Omega}_{IMU}$ because the body frame $< B >$ was taken to be coincident to the IMU frame.

$$\boldsymbol{\zeta}_{IMU} = \begin{bmatrix} a_x \\ \Omega_z \\ \Omega_{DC} \end{bmatrix} \approx \begin{bmatrix} a_1 V + b_1 u_1 + g \sin\beta_{IMU} \\ \frac{\tan\phi}{L} V \\ \frac{k_g}{2\pi r} V \end{bmatrix} = \boldsymbol{h}_{IMU}(\boldsymbol{\xi}), \tag{4.10}$$

where $g(= 9.81\,\frac{m}{s^2})$ is the gravity acceleration, $r(= 14\,mm)$ is the radius of the robot's wheels, and $k_g(= 3.4)$ is the transmission gear ratio. The angle β_{IMU} gives the rotation about the y-axis between the body frame $< B >$ and the inertial frame $< I >$ and was computed by integrating the following equations:[4]

$$\dot{\boldsymbol{\Theta}}_{IMU} = \begin{bmatrix} \dot{\alpha}_{IMU} \\ \dot{\beta}_{IMU} \\ \dot{\gamma}_{IMU} \end{bmatrix} = \begin{bmatrix} 1 & \sin\alpha_{IMU}\tan\beta_{IMU} & \cos\alpha_{IMU}\tan\beta_{IMU} \\ 0 & \cos\alpha_{IMU} & -\sin\alpha_{IMU} \\ 0 & \frac{\sin\alpha_{IMU}}{\cos\beta_{IMU}} & \frac{\cos\alpha_{IMU}}{\cos\beta_{IMU}} \end{bmatrix} \boldsymbol{\Omega}_{IMU}. \tag{4.11}$$

The initial condition of the integrator on $\boldsymbol{\Theta}$ was taken as $\boldsymbol{\Theta}_0 = [atan(\bar{a}_{y0}, \bar{a}_{z0})\; asin(\frac{\bar{a}_{x0}}{g})\; 0]^T$, where $\bar{a}_{x0}, \bar{a}_{y0}, \bar{a}_{z0}$ are the mean values of the accelerometer's measurements along the three axes delivered within $1\,s$ before starting each test while the robot was not moving. The covariance matrix of the model uncertainties was the same in both the OF-based and IMU-based EKF, i.e. $Q_{IMU} = Q$, whereas the covariance matrix of the measurement noise was identified as $R_{IMU} = diag(0.263^2, 0.047^2, 0.001^2)$.

The ground-truth values were computed from the robot's pose measurements after being low-pass filtered at $f_{c_l} = 100\,Hz$ ($[X'\,Y'\,Z'\,\alpha'\,\beta'\,\gamma']^T$) as follows (see section 4.2.4 for details):

- $V_{truth} = \sqrt{\dot{X'}^2 + \dot{Y'}^2}$;
- $\phi_{truth} = \arctan(L\frac{\dot{\gamma'}}{V_{truth}})$;
- $h^l_{truth} = h_0 + Z'' + y_l \sin(\alpha'')\cos(\beta'') + z_l \cos(\alpha'')\cos(\beta'')$ (same for h^r_{truth} using y_r, z_r);
- $\omega^l_{i_{truth}} = \frac{(y_l\dot{\gamma'} - V_{truth})\sin^2\varphi_i}{h^l_{truth}}$ (same for $\omega^r_{i_{truth}}$ using $y_r,\ h^r_{truth}$);

where h_0 is the height of the body frame $< B >$ with respect to the local ground plane, i.e. $h_0 = h - z_{l,r} = 50\,mm$, and Z'', α'', β'' are the high-pass filtered values of Z', α', β' ($f_{c_h} = 1\,Hz$), which were adopted in order to cut off the low-frequency components due to the changes in the ground's slope and height.

The average refresh rate of the sensors $\bar{f}_\omega$ was computed by dividing the number of median values of the OF measurements obtained during each test by the time taken to run the test. The precision of the median values of the OF measurements $\omega^l_m,\ \omega^r_m$ as well as that of the estimates $\hat{V}, \hat{\phi}$ was computed by dividing the standard deviations of the errors with respect to their ground-truth values by their average absolute value (e.g. if $|\bar{V}| = 0.3\,\frac{m}{s}$ and $\sigma_V = 0.007\,\frac{m}{s}$ then the precision will be $\frac{0.007}{0.3} \approx 2.3\%$). The accuracy of the estimates of the robot's absolute position $\hat{\boldsymbol{X}}$ and orientation $\hat{\theta}$, instead, was computed by

[4] Note that we substituted $\dot{\gamma}_{IMU}$ with Ω_z in the third equation in (4.10) because in the experiments presented here we always had $\Omega_z \approx \dot{\gamma}_{IMU}$. Also note that the angle β_{IMU} was considered as a parameter although it depends on Ω_z, and therefore on the state $\boldsymbol{\xi}$, since its value varied within a small range and no significant improvement was obtained using a more complex output model.

dividing the maximum errors with respect to the ground-truth position and orientation respectively by the distance traveled and the angle covered when these maxima were reached (e.g. if $||\boldsymbol{X} - \hat{\boldsymbol{X}}|| = 0.7\,m$ after traveling $45\,m$ then the accuracy will be $\frac{0.7}{45} \approx 1.6\%$). We preferred to use this traditional metrics for the odometry accuracy instead of using other common metrics in visual odometry based on the errors in the relative camera poses, such as that presented in [311], because such metrics give values that are well representative of the "noise" in the position and orientation estimates (error between contiguous estimates), but not necessarily of the accuracy of these estimates (absolute error). To better evaluate the odometry performances during the entire path, we also provide here the plots of the error in the position and orientation with respect to time and distance.

Indoors

Figure 4.6 shows the results obtained indoors when the robot was driven on the floor shown in Fig. 4.1(c) on a circular path, keeping a constant steering angle at a velocity ranging from 0.3 to $1.3\,\frac{m}{s}$, under dynamically changing lighting conditions. The OF resolution $\Delta\omega^*$ required was set at $0.3\,\frac{rad}{s}$ and the average refresh rate $\bar{f}_\omega$ obtained was about $327\,Hz$.

First the incoming sunlight was made to vary by up to 2 decades by slowly closing the 8 blinds (from 0 to $24\,s$), giving a maximum luminosity of about $5000\,Lux$ ($I_{ph} \approx 5 \times 10^{-5}\,A$), then the artificial lighting was made to vary both slowly and rapidly by about 2 decades, by varying the neon ceiling lights and then switching them off (from 24 to $46\,s$). The robot was then driven with only the Vicon cameras' LED lights switched on (from 46 to $59\,s$), corresponding to a luminosity of about $0.3\,Lux$ (average I_{ph} of about $3 \times 10^{-10}\,A$), before switching the neon lights on again. The light levels tested therefore covered a nearly 6-decade range (I_{ph} from about 10^{-10} to $5 \times 10^{-5}\,A$), as can be seen from Fig. 4.6(f).

Indoor tests at constant light levels were also carried out in the same environment and using the same trajectory to determine the highest precision and widest range achievable with our novel OF sensors. First the robot was driven at constant velocities from the minimum to the maximum values, i.e. from 0.3 to $1.5\,\frac{m}{s}$, setting the OF resolution $\Delta\omega^*$ required at $0.05\,\frac{rad}{s}$: Fig. 4.7(A) shows the case where the robot was driven at its maximum velocity. The minimum OF resolution $\Delta\omega^*$ used was set at $0.05\,\frac{rad}{s}$ in order to obtain a measurable OF range of $1.5\,\frac{rad}{s}$ without causing saturation of the CPU load. The velocity was then varied linearly from the minimum to the maximum values with accelerations varying from ± 0.3 to $\pm 1\,\frac{m}{s^2}$ (Fig. 4.7(B)). In the latter case, the height of the VMSs with respect to the ground h was exceptionally decreased to $135\,mm$ and the OF resolution $\Delta\omega^*$ required was increased to $0.5\,\frac{rad}{s}$ in order to achieve a larger OF range (from 1.5 to $15\,\frac{rad}{s}$). The average refresh rate $\bar{f}_\omega$ obtained was about $333\,Hz$ in all the cases tested.

Lastly, indoor tests were carried out using various trajectories and various floor textures. Figure 4.8 shows the results obtained indoors when the robot was driven on paths forming a figure of eight and a square, while the velocity and the steering angle were made to vary trapezoidally. The OF resolution $\Delta\omega^*$ required was set at 0.1 and $0.2\,\frac{rad}{s}$ and the average refresh rate $\bar{f}_\omega$ obtained was about $325\,Hz$.

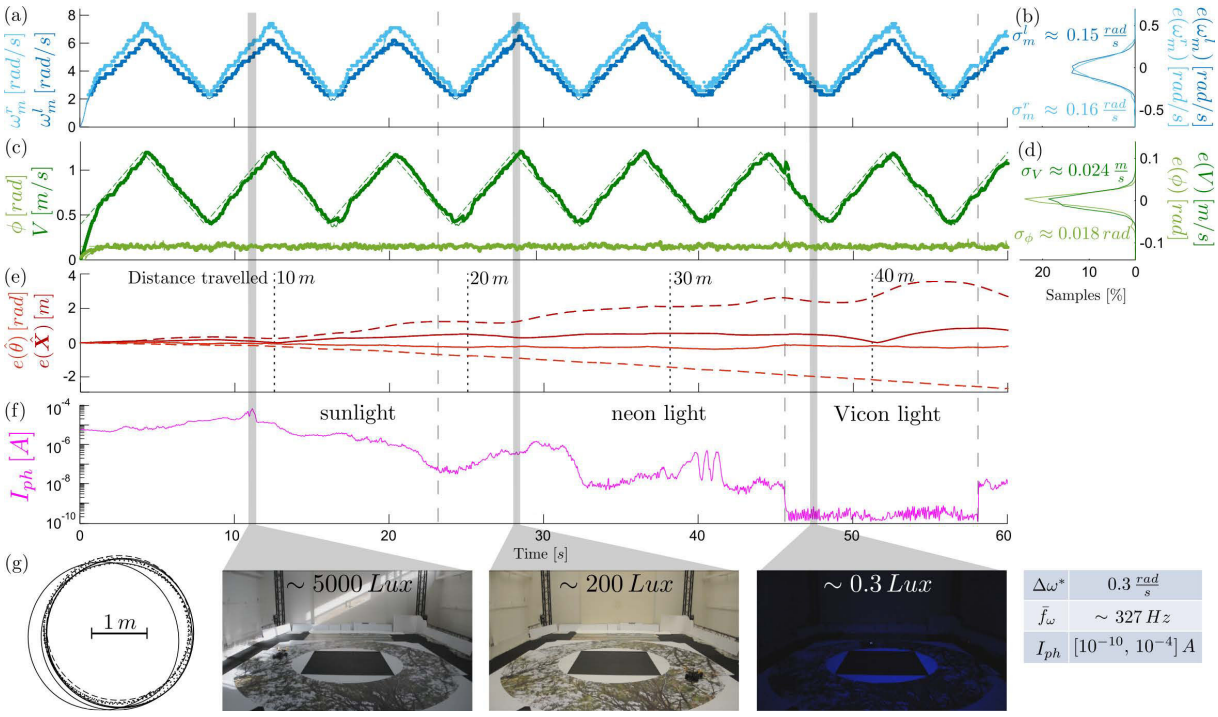

Figure 4.6: OF measurements and odometry results robust to changes in the light levels in a nearly 6-decade range obtained when the robot was driven on a circular path at a velocity ranging from 0.3 to 1.3 $\frac{m}{s}$. (a) Median values ω_m^l, ω_m^r (dots) and ground-truth values $\omega_{m_{truth}}^l, \omega_{m_{truth}}^r$ (solid lines) of the OF produced on the left (dark blue) and right (light blue) sides. (b) Distribution of the errors between the median values and the ground-truth values shown in Fig. 4.6(a), i.e. $\omega_{m_{truth}}^l - \omega_m^l$ (dark blue line) and $\omega_{m_{truth}}^r - \omega_m^r$ (light blue line). (c) EKF estimates $\hat{V}, \hat{\phi}$ (dots), ground-truth values V_{truth}, ϕ_{truth} (solid lines) and reference values V^*, ϕ^* (dashed lines) of the robot's longitudinal velocity (dark green) and steering angle (light green). (d) Distribution of the errors between the EKF estimates and the ground-truth values shown in Fig. 4.6(c), i.e. $V_{truth} - \hat{V}$ (dark green line) and $\phi_{truth} - \hat{\phi}$ (light green line). (e) Errors in the robot's position and orientation estimated using $\hat{V}, \hat{\phi}$, i.e. $||\boldsymbol{X} - \hat{\boldsymbol{X}}||$ (dark red solid line) and $\theta - \hat{\theta}$ (light red solid line), and $\hat{V}_{IMU}, \hat{\phi}_{IMU}$, i.e. $||\boldsymbol{X} - \hat{\boldsymbol{X}}_{IMU}||$ (dark red dashed line) and $\theta - \hat{\theta}_{IMU}$ (light red dashed line). (f) Photo-current response of the photodiode placed next to the LMS on the left side. The current values were computed taking $I_{Ph} = I_{dark}(e^{8.8V_{out}} - 1)$, where $I_{dark}(\approx 0.1\,nA)$ is the dark current and V_{out} is the amplifier's output voltage. (g) Reference $\boldsymbol{X}^*$ (dashed line), estimated $\hat{\boldsymbol{X}}$ (dotted line) and ground-truth $\boldsymbol{X}$ (solid line) robot's trajectory in the closed-loop mode. The pictures at the bottom of the figure show, from left to right, the test environment under incoming sunlight, neon light and the Vicon cameras' LED light. The light values in *Lux* were estimated using the current-to-illuminance curve given in the photodiode's datasheet.

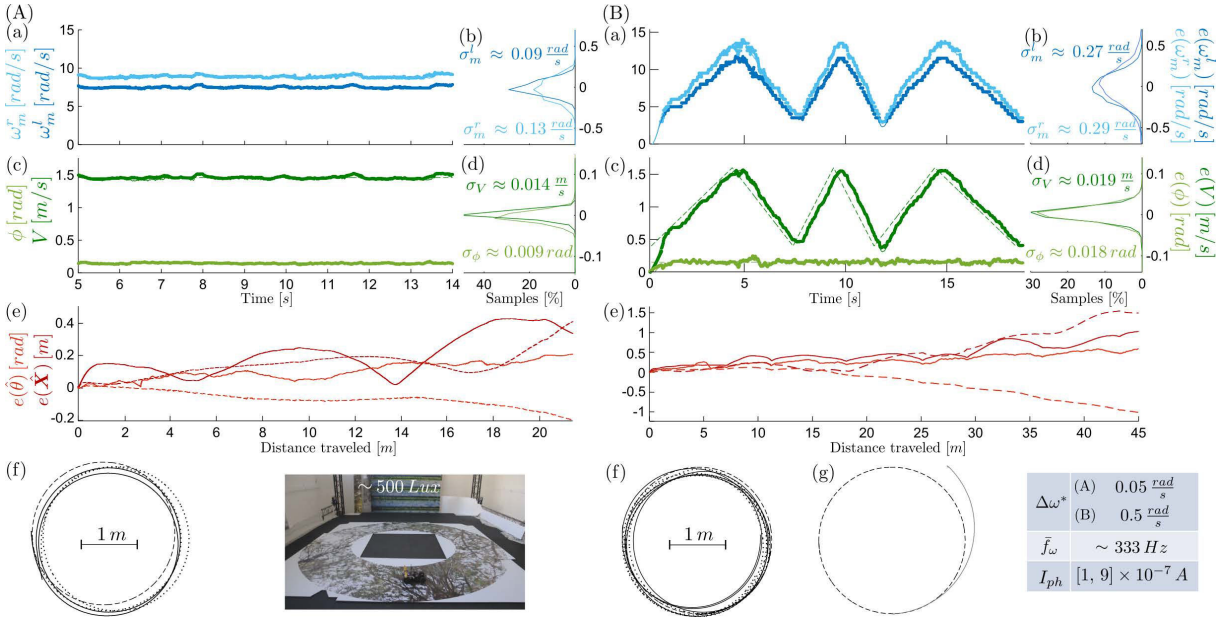

Figure 4.7: Robust OF measurements and odometry results obtained (A) at high resolution and (B) in a wide OF range, when the robot was driven on a circular path (A) at its maximum velocity (1.5 $\frac{m}{s}$) and (B) at a velocity ranging from 0.3 to 1.5 $\frac{m}{s}$ with various accelerations. (a) ω_m^l, ω_m^r (dots) and $\omega_{m_{truth}}^l, \omega_{m_{truth}}^r$ (solid lines) during 1 lap. (b) Distribution of $\omega_{m_{truth}}^l - \omega_m^l$ (dark blue line) and $\omega_{m_{truth}}^r - \omega_m^r$ (light blue line) during 3-5 laps. (c) $\hat{V}, \hat{\phi}$ (dots), V_{truth}, ϕ_{truth} (solid lines) and V^*, ϕ^* (dashed lines) during 1 lap. (d) Distribution of $V_{truth} - \hat{V}$ (dark green line) and $\phi_{truth} - \hat{\phi}$ (light green line) during 3-5 laps. (e) $||\boldsymbol{X} - \hat{\boldsymbol{X}}||$, $\theta - \hat{\theta}$ (solid lines) and $||\boldsymbol{X} - \hat{\boldsymbol{X}}_{IMU}||$, $\theta - \hat{\theta}_{IMU}$ (dashed lines), during 3-5 laps. (f) $\boldsymbol{X}^*$ (dashed line), $\hat{\boldsymbol{X}}$ (dotted line) and $\boldsymbol{X}$ (solid line) in the closed-loop mode. (g) $\boldsymbol{X}^*$ (dashed line) and $\boldsymbol{X}$ (gray line) in the open-loop mode.

The results presented here show that the OF measurements did not depend on either the average light level or the changes in the light, except in the case of very large, sudden changes (see Fig. 4.6(a) at about 46 s and 59 s), nor they depended on the visual patterns and the types of trajectory used (no statistical evidence in the error distributions and in the refresh frequencies). In addition, although the error in the OF measurements increased slightly with the OF magnitude, our novel sensors were able to measure the OF in a very large range, from about 1.5 to 15 $\frac{rad}{s}$ (i.e. from about 85 to 850 $\frac{^o}{s}$), and with relatively high OF resolution (0.05 $\frac{rad}{s}$). However, both a high resolution, and therefore precision, and a wide range could not be provided yet at the same time due to computational limitations.

In all these tests, the overall error distributions were nearly Gaussian with a quasi-zero mean (Fig. 4.6(b), 4.7(A)(b),(B)(b), 4.8(A)(b),(B)(b)), giving values of the adjusted R-square goodness-of-fit statistics, computed using the Matlab Curve Fitting Toolbox, always greater than 0.99. However, it is worth noting that the error distributions were sometimes more peaky than their fitted Gaussian curves, meaning that the errors were statistically closer to their mean values than those obtained with a pure Gaussian

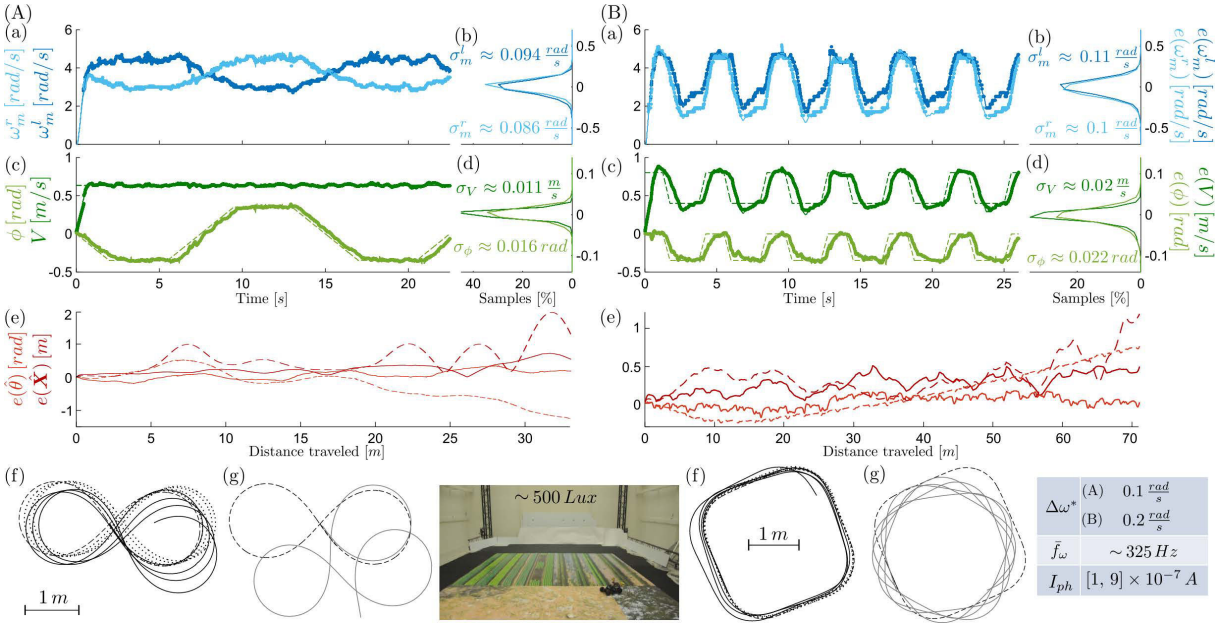

Figure 4.8: Robust OF measurements and odometry results obtained when the robot was driven on a path forming (A) a figure of eight and (B) a square. (a) ω_m^l, ω_m^r (dots) and $\omega_{m_{truth}}^l, \omega_{m_{truth}}^r$ (solid lines) during 1.5 laps. (b) Distribution of $\omega_{m_{truth}}^l - \omega_m^l$ (dark blue line) and $\omega_{m_{truth}}^r - \omega_m^r$ (light blue line) during 3-8 laps. (c) $\hat{V}, \hat{\phi}$ (dots), V_{truth}, ϕ_{truth} (solid lines) and V^*, ϕ^* (dashed lines) during 1.5 laps. (d) Distribution of $V_{truth} - \hat{V}$ (dark green line) and $\phi_{truth} - \hat{\phi}$ (light green line) during 3-8 laps. (e) $||\boldsymbol{X} - \hat{\boldsymbol{X}}||$, $\theta - \hat{\theta}$ (solid lines) and $||\boldsymbol{X} - \hat{\boldsymbol{X}}_{IMU}||$, $\theta - \hat{\theta}_{IMU}$ (dashed lines), during 3-8 laps. (f) $\boldsymbol{X}^*$ (dashed line), $\hat{\boldsymbol{X}}$ (dotted line) and $\boldsymbol{X}$ (solid line) in the closed-loop mode. (g) $\boldsymbol{X}^*$ (dashed line) and $\boldsymbol{X}$ (gray line) in the open-loop mode.

distribution. We also note that such a difference in the height and shape of the error distributions was due to the fact that these distributions were extracted from histograms over a rather low-resolution grid of beams ($\Delta beam$ from 0.02 to 0.04) in order to have smooth curves.

The greatest OF precision obtained was about 1.2% when $\Delta\omega^*$ was set at $0.05\ \frac{rad}{s}$, whereas the lowest was about 3.6% when $\Delta\omega^*$ was set at $0.5\ \frac{rad}{s}$. The robot adopted the required velocity and steering angle, giving a precision of the robot's velocity and steering angle estimates ($\hat{V}, \hat{\phi}$) ranging from 1% to 5% and from 6% to 12%, respectively, which made it possible to drive the robot relatively close to the reference trajectory (Fig. 4.6(g), 4.7(A)(f),(B)(f), 4.8(A)(f),(B)(f)). When the robot was driven in the open-loop mode, however, i.e. when setting $\bar{\boldsymbol{\xi}} = \boldsymbol{\xi}^*$ instead of applying the control scheme presented in Fig. 4.5, the position and orientation errors were much larger than those obtained in the closed-loop mode (Fig. 4.7(B)(g), 4.8(A)(g),(B)(g)). Depending on the type of trajectory and on whether the robot was under or over-steered with respect to the steering values required, the robot could go very astray, sometimes making the robot drift up to going off the carpet at a very early stage (Fig. 4.7(B)(g), 4.8(A)(g)), or stay in a closed path although very different to that required (Fig. 4.8(B)(g)).

By using the OF-based estimates of the robot's velocity and steering angle for odometry purposes, we obtained an accuracy of the position and orientation estimates ($\hat{\boldsymbol{X}}$, $\hat{\theta}$) ranging from 0.3% to 2.3%, after the robot had been traveling from 18 to $71\,m$ and turning from 4.1π to $8.1\pi\,rad$, i.e. after about 2 to 8 laps. It is worth noting that such an odometry accuracy was about 2 to 9 times higher than that obtained by using the IMU-based estimates of the robot's velocity and steering angle ($\hat{V}_{IMU}, \hat{\phi}_{IMU}$) (Fig. 4.6(e), 4.7(B)(e), 4.8(A)(e),(B)(e)), except in the case of constant high velocity where the odometry performances were comparable (Fig. 4.7(A)(e)). Such a result could be due to the fact that the OF-based estimates ($\hat{V}, \hat{\phi}$) were not very accurate at the very beginning of the test because no measurements were delivered until the OF values reached the range required, i.e. $[6.75, 8.25]\,\frac{rad}{s}$ and $[8.25, 9.75]\,\frac{rad}{s}$ for the left and right side, respectively, giving large errors in the robot's position and orientation estimates (see the first part of Fig. 4.7(A)(e)).

Some quantitative data and statistics on each indoor test are given in Table 5.1 in the appendix.

Outdoors

Figures 4.9, 4.10 show the results obtained outdoors during the day-time (at about 12:00 PM) and at night (at about 20:00 PM), respectively, when the robot was driven on the ground shown in Fig. 4.1(d) on a circular path, keeping a constant steering angle while varying the velocity from 0.3 to $1.3\,\frac{m}{s}$. The OF resolution $\Delta\omega^*$ required was set at $0.3\,\frac{rad}{s}$ and the average refresh rate $\bar{f}_\omega$ obtained was about $300\,Hz$ in both tests.

Outdoor tests were also carried out in the same environment using various ground textures (Fig. 4.11) and various trajectories (Fig. 4.12). In the first case, 4 regions with different textures were added to the floor: (1) a mixture of grass, leaves and flowers, (2) dark sand and (3) light sand, and (4) black asphalt gravel (see the pictures at the bottom of Fig. 4.11). The OF resolution $\Delta\omega^*$ required was set at 0.2 and $0.1\,\frac{rad}{s}$ and the average refresh rate $\bar{f}_\omega$ obtained in the 2 tests was about 295 and $328\,Hz$, respectively.

The results presented here show that our sensors responded appropriately outdoors by delivering OF measurements regardless of the luminosity in a 6-decade range (I_{ph} ranging from about 10^{-9} to $10^{-3}\,A$) and despite the shadows produced by the robot itself under daylight conditions (see, for instance, pictures (1) and (2) in Fig. 4.9). In addition, our sensors delivered OF measurements that did not depend significantly on either the trajectory or the ground textures (grass, leaves, flowers, dark and light sand, dark and light gravel), although sometimes fewer measurements were delivered on the left side when the robot was driving on some of the textured regions (Fig. 4.11(a)).

As in the indoor tests, the overall error distributions were consistently nearly Gaussian with a quasi-zero mean (Fig. 4.9(b), 4.10(b), 4.11(b), 4.12(b)), giving values of the adjusted R-square goodness-of-fit statistics, computed using the Matlab Curve Fitting Toolbox, always greater than 0.98. The same observations on the height and shape of the error distributions done for the indoor tests can also be done for the outdoor tests.

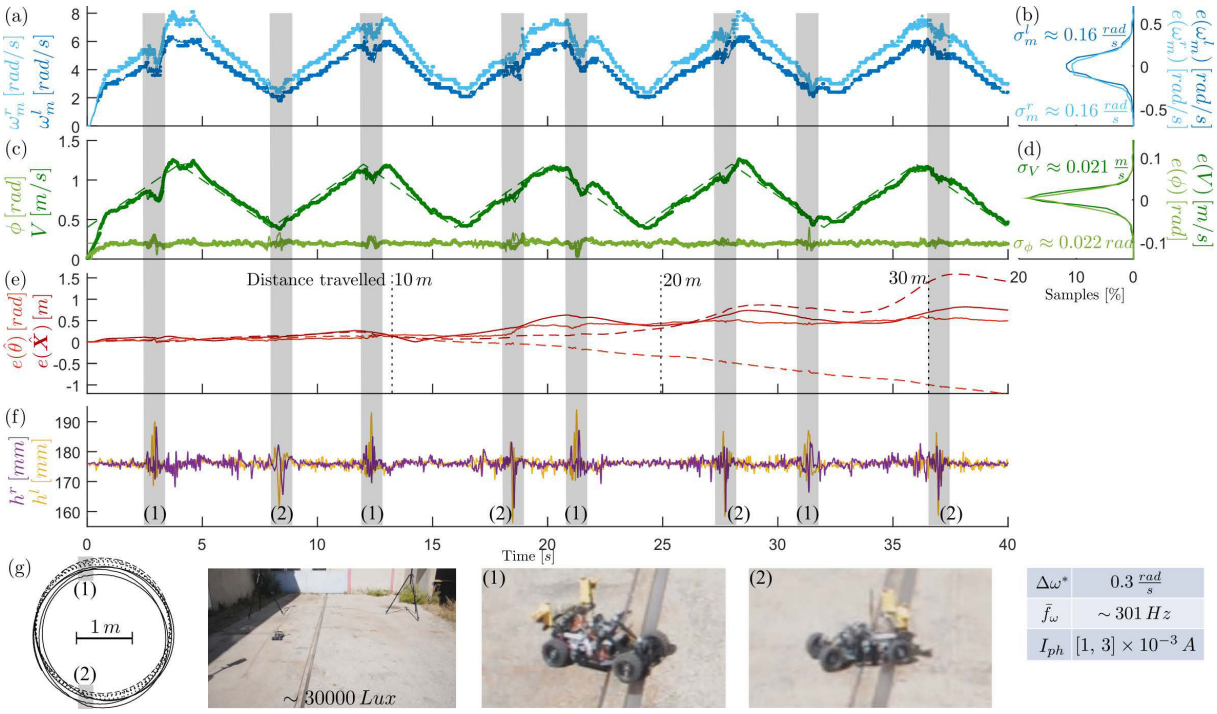

Figure 4.9: Robust OF measurements and odometry results obtained outdoors at a very high light level (at about 12:00 PM) when the robot was driven on a circular path at a velocity ranging from 0.3 to 1.3 $\frac{m}{s}$. (a) ω_m^l, ω_m^r (dots) and $\omega_{m_{truth}}^l, \omega_{m_{truth}}^r$ (solid lines). (b) Distribution of $\omega_{m_{truth}}^l - \omega_m^l$ (dark blue line) and $\omega_{m_{truth}}^r - \omega_m^r$ (light blue line). (c) $\hat{V}, \hat{\phi}$ (dots), V_{truth}, ϕ_{truth} (solid lines) and V^*, ϕ^* (dashed lines). (d) Distribution of $V_{truth} - \hat{V}$ (dark green line) and $\phi_{truth} - \hat{\phi}$ (light green line). (e) $||\boldsymbol{X} - \hat{\boldsymbol{X}}||$, $\theta - \hat{\theta}$ (solid lines) and $||\boldsymbol{X} - \hat{\boldsymbol{X}}_{IMU}||$, $\theta - \hat{\theta}_{IMU}$ (dashed lines). (f) Left (yellow line) and right (purple line) local sensor's height h^l, h^r approximately estimated using the robot's pose measurements. (g) $\boldsymbol{X}^*$ (dashed line), $\hat{\boldsymbol{X}}$ (dotted line) and $\boldsymbol{X}$ (solid line). The picture in the bottom-left part of the figure shows the test environment at about 12:00 PM, when the light level was about 30000 Lux. The pictures marked (1) and (2) show 2 situations where the robot was driving on the steel rail, corresponding to the shaded regions in Fig. 4.9(a-g).

The OF precision obtained here was slightly higher than that obtained in the indoor tests, ranging from 2.6% to 3.8%, and did not significantly depend on the light levels, the ground textures or the type of trajectory.[5] However, the average precision of the robot's velocity and steering angle estimates $(\hat{V}, \hat{\phi})$ is similar to that obtained in the indoor tests (2.3% vs 2.4% and 9.3% vs 9.8%, respectively), showing the robustness of the method presented here to vibrations liable to cause a change in the sensors' local height of up to $\pm 10\,mm$ (Fig. 4.9(f), 4.10(f)), amounting to about 6 % of the nominal height. In the case

[5] The standard deviations and the corresponding precision in the first 2 tests (Fig. 4.9, 4.10) were computed without taking into account the moments when the robot was driving on the steel rail (shaded regions in Fig. 4.9(a-g), 4.10(a-g)) as the impact with the rail produced very large vibrations.

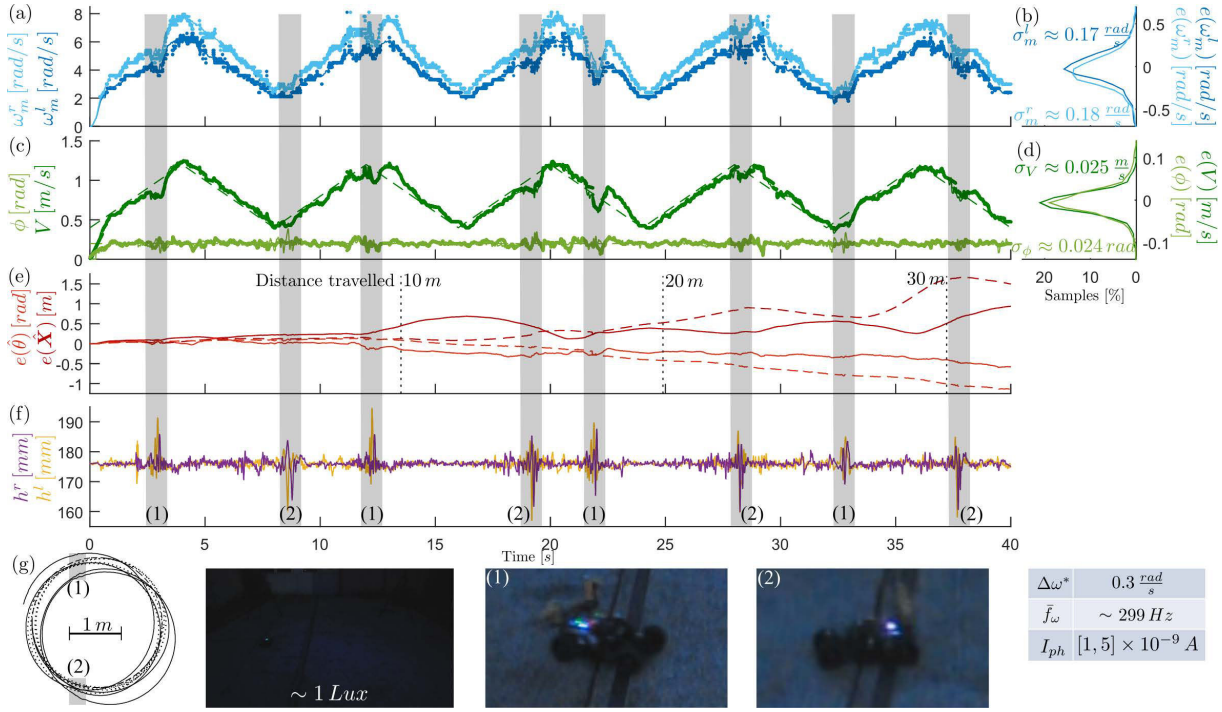

Figure 4.10: Robust OF measurements and odometry results obtained outdoors at very low light levels (at about 20:00 PM) when the robot was driven on a circular path at a velocity ranging from 0.3 to 1.3 $\frac{m}{s}$. (a) ω_m^l, ω_m^r (dots) and $\omega_{m_{truth}}^l, \omega_{m_{truth}}^r$ (solid lines). (b) Distribution of $\omega_{m_{truth}}^l - \omega_m^l$ (dark blue line) and $\omega_{m_{truth}}^r - \omega_m^r$ (light blue line). (c) $\hat{V}, \hat{\phi}$ (dots), V_{truth}, ϕ_{truth} (solid lines) and V^*, ϕ^* (dashed lines). (d) Distribution of $V_{truth} - \hat{V}$ (dark green line) and $\phi_{truth} - \hat{\phi}$ (light green line). (e) $||\boldsymbol{X} - \hat{\boldsymbol{X}}||$, $\theta - \hat{\theta}$ (solid lines) and $||\boldsymbol{X} - \hat{\boldsymbol{X}}_{IMU}||$, $\theta - \hat{\theta}_{IMU}$ (dashed lines). (f) h^l (yellow line) and h^r (purple line). (g) $\boldsymbol{X}^*$ (dashed line), $\hat{\boldsymbol{X}}$ (dotted line) and $\boldsymbol{X}$ (solid line). The picture in the bottom-left part of the figure shows the test environment at about 20:00 PM, when the light level was about $1\,Lux$. The images marked (1) and (2) show 2 situations where the robot was driving on the steel rail, corresponding to the shaded regions in Fig. 4.10(a-g).

of larger changes in height, such as those caused by the rail (shaded regions in Fig. 4.9(a-g), 4.10(a-g)), the robot could still estimate V and ϕ but with a much lower precision, and a more complex method of estimation including information from other sensors would therefore have to be implemented in order to achieve a higher level of robustness. In fact, the error in the position and orientation estimates ($\hat{\boldsymbol{X}}$, $\hat{\theta}$) seems to particularly increase when the robot was driving on the rail at relatively high speed, because of larger errors in the OF-based velocity and steering angle estimates (see, for instance, the shaded regions in Fig. 4.9 at about $18\,s$ and in Fig. 4.10 at about $12\,s$). However, even in presence of such large vibrations, we obtained a still good accuracy of the position and orientation estimates, i.e. from 2% to 3% after the robot had been traveling about $30\,m$ and turning about 6-$8\pi\,rad$ (i.e. about 3 to 4 laps), which was slightly higher than that obtained by using the IMU-based estimates of the robot's velocity and steering

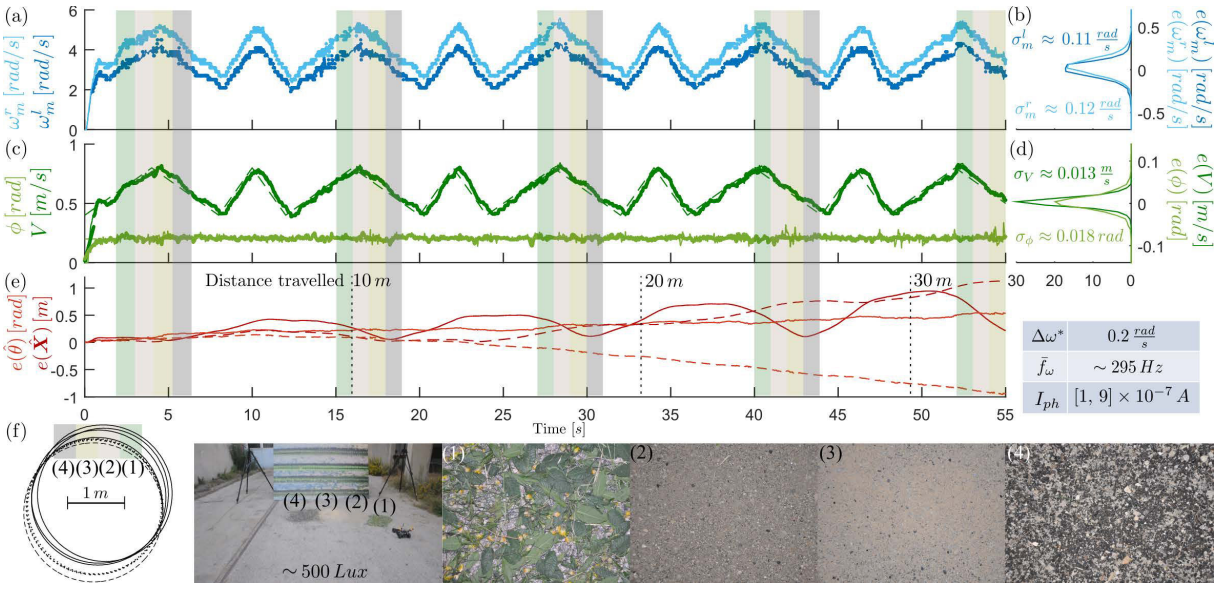

Figure 4.11: Robust OF measurements and odometry results obtained outdoors on various ground textures. The robot was driven on a circular path at a velocity ranging from 0.4 to 0.8 $\frac{m}{s}$ on the ground shown in Fig. 4.1(d), to which 4 areas with different textures were added. (a) ω_m^l, ω_m^r (dots) and $\omega_{m_{truth}}^l, \omega_{m_{truth}}^r$ (solid lines). (b) Distribution of $\omega_{m_{truth}}^l - \omega_m^l$ (dark blue line) and $\omega_{m_{truth}}^r - \omega_m^r$ (light blue line). (c) $\hat{V}, \hat{\phi}$ (dots), V_{truth}, ϕ_{truth} (solid lines) and V^*, ϕ^* (dashed lines). (d) Distribution of $V_{truth} - \hat{V}$ (dark green line) and $\phi_{truth} - \hat{\phi}$ (light green line). (e) $||\boldsymbol{X} - \hat{\boldsymbol{X}}||$, $\theta - \hat{\theta}$ (solid lines) and $||\boldsymbol{X} - \hat{\boldsymbol{X}}_{IMU}||$, $\theta - \hat{\theta}_{IMU}$ (dashed lines). (f) $\boldsymbol{X}^*$ (dashed line), $\hat{\boldsymbol{X}}$ (dotted line) and $\boldsymbol{X}$ (solid line). The pictures at the bottom of the figure show (from left to right) the test environment and the 4 textured regions: (1) a mixture of grass, leaves and flowers, (2) dark sand and (3) light sand, and (4) black asphalt gravel, corresponding to the shaded regions in Fig. 4.11(a-f).

angle ($\hat{V}_{IMU}, \hat{\phi}_{IMU}$) (Fig. 4.9(e), 4.10(e), 4.11(e), 4.12(e)). In particular, the OF-based odometry gave better performances in average than the IMU-based odometry when using trajectories where the steering angle, and therefore the angular velocity, changed in sign, i.e. when driving on an eight-shaped path (Fig. 4.12(e)). However, we note that, in the outdoor tests, the errors in the IMU-based estimates were often lower than those in the OF-based estimates in the first part of the tests and then started increasing very rapidly, probably because of the gyroscope's varying bias. All in all, we note that the accuracy of the solely IMU-based odometry strongly depends on the non-linear dynamics intrinsic to the IMU used, therefore better odometry results could be obtained by including these non-linearities in the EKF.

Some quantitative data and statistics on each outdoor test are given in Table 5.2 in the appendix.

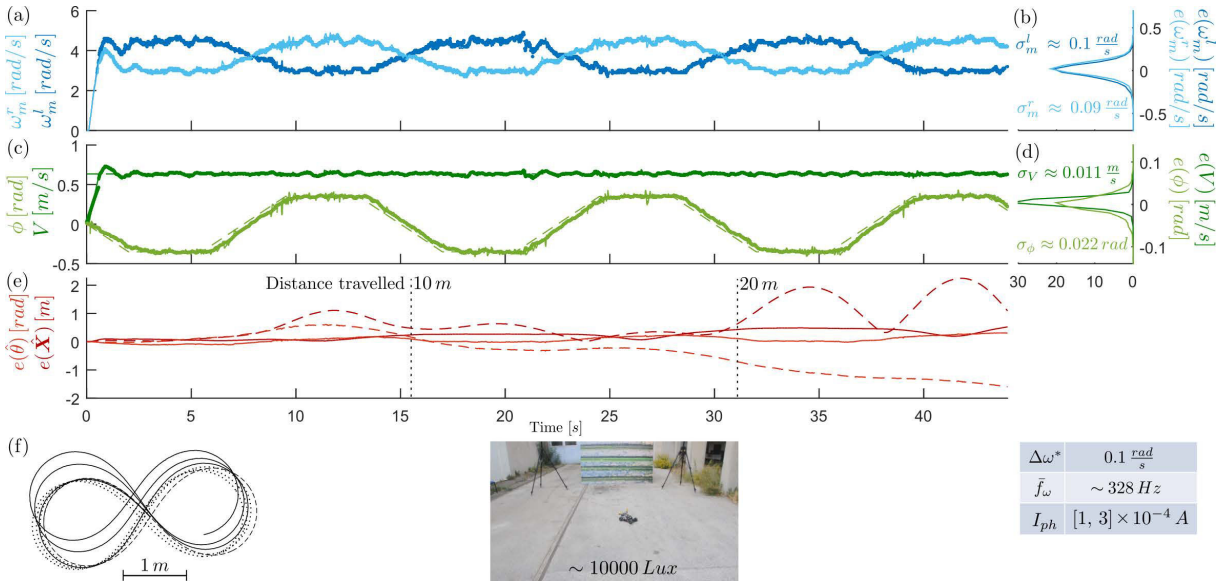

Figure 4.12: Robust OF measurements and odometry results obtained outdoors when the robot was driven on a path forming a figure of eight at a constant velocity while the steering angle varied trapezoidally. (a) ω^l_m, ω^r_m (dots) and $\omega^l_{m_{truth}}, \omega^r_{m_{truth}}$ (solid lines). (b) Distribution of $\omega^l_{m_{truth}} - \omega^l_m$ (dark blue line) and $\omega^r_{m_{truth}} - \omega^r_m$ (light blue line). (c) $\hat{V}, \hat{\phi}$ (dots), V_{truth}, ϕ_{truth} (solid lines) and V^*, ϕ^* (dashed lines). (d) Distribution of $V_{truth} - \hat{V}$ (dark green line) and $\phi_{truth} - \hat{\phi}$ (light green line). (e) $||\boldsymbol{X} - \hat{\boldsymbol{X}}||$, $\theta - \hat{\theta}$ (solid lines) and $||\boldsymbol{X} - \hat{\boldsymbol{X}}_{IMU}||$, $\theta - \hat{\theta}_{IMU}$ (dashed lines). (f) $\boldsymbol{X}^*$ (dashed line), $\hat{\boldsymbol{X}}$ (dotted line) and $\boldsymbol{X}$ (solid line). Note that sometimes the noise in the measurements delivered by the Vicon system was higher than the actual measurement error because the markers were poorly detected (see, for instance, the green solid line in Fig. 4.12(c)).

4.2.6 Conclusions

The low-cost car-like robot called BioCarBot presented in this paper is able to estimate its velocity and steering angle, and therefore its position and orientation via an Extended Kalman Filter (EKF), using only the OF measurements delivered by two novel downward-facing VMSs. Thanks to the cross-correlation method and the auto-adaptive pixels used, these novel VMSs have the following advantages:

- the OF measurements are robust to high-dynamic-range lighting conditions (in a 7-decade range with sharp changes of up to 2 decades within $0.5\,s$) and to the various visual patterns encountered;
- the refresh rate of the OF measurements is relatively high ($300 - 333\,Hz$) and nearly constant, and does not depend on the bandwidth of the band-pass filter;
- the resolution on the OF measurements is also relatively high (up to $0.05\,\frac{rad}{s}$) and constant, i.e. it does not depend on the OF magnitude, and can be set at whatever value required;
- the OF measurement range is relatively wide (from 1.5 to $15\,\frac{rad}{s}$, i.e. 85 to $850\,\frac{^o}{s}$) and can also be

adjusted as required;

- the sensor's precision can be relatively high (3.8% in the worst case), depending on the OF resolution required, the OF range and the robot's vibrations.

With these OF sensors, the robot was able to estimate its own velocity and steering angle, and therefore its position and orientation accurately, both indoors and outdoors, and the drifts liable to occur when it was driven in the open-loop mode were greatly reduced. The main features of the minimalistic visual odometry method presented here can be listed as follows:

- robustness to high-range light levels (by daylight and at night, i.e. from about 1 to $30000\,Lux$, including shadows) and various ground textures (asphalt, gravel, sand, leaves, etc.);
- robustness to vibrations liable to affect change in the sensors' local height up to 6% (i.e. from $\pm 10\,mm$ to $175\,mm$);
- fairly high-accuracy position and orientation estimation (from 0.3% to 3%) after covering a distance from $30\,m$ to $75\,m$ and turning through an angle from 6π to $16\pi\,rad$.

We also showed that the accuracy was higher in average than that of the IMU-based odometry, especially in the indoor tests and when the distance and the rotation angle covered were relatively large.

It is worth noting that both the sampling frequency, and hence the refresh rate, and the OF resolution, and hence the precision, could be increased by either optimizing the implementation of the algorithm or increasing the computational resources. These 2 parameters could be also adjusted in real time to the values required depending on the robot's reference trajectory, and therefore on the OF setpoint profile.

The precision of the estimates could also be improved by (i) adapting the EKF to include just a single OF measurement (i.e. on one side only) when no measurement is available on the other side, and (ii) precisely adjusting the values of the measurement and process covariance matrices of the EKF after performing a calibration phase prior to each test. In addition, the OF and inertial measurements could be combined together in one EKF by including accurately-identified dynamic models of the sensors used which take into account their intrinsic non-linearities, in order to achieve greater robustness and precision, as required for the autonomous vehicles of the future [312] and even tomorrow's flying robots [290].

Tests involving more complex and longer trajectories as well as more challenging environments, such as forest paths, are now being considered, and we are planning to test the sensors and the estimation method presented here on a real vehicle with a view to improving the existing odometry techniques. In this view, we are willing to apply the improvements discussed above and to evaluate the results obtained also by using the metrics presented in [311].

4.3 Conclusion

In the paper presented here, we showed that the new optic flow sensor based on the M^{2}APix pixels and the improved version of the correlation-based time-of-travel method delivered robust measurements in real road conditions. In particular, this LMS provides some advantages with respect to those previously developed at our laboratory (see section 1.8.4):

- the refresh rate and the precision of the measurements nearly did not depend on the operating conditions (visual patterns, light levels, vibrations, etc), which often occurred when using a thresholding phase;
- the refresh rate was higher ($\sim 300\,Hz$) and did not depend on the band-pass filtering phase;
- the measurable OF range was wider (85 to $850\,\frac{^o}{s}$);
- the OF resolution was constant (i.e. did not depend on the OF amplitude itself) and can theoretically be set as whatever value as required.

However, the method proposed has the main drawback of being very consuming in terms of computational cost, especially if we want both wide range and high resolution, and therefore should be implemented on embedded electronics allowing parallel computing, such as DSPs or FPGAs.

We also showed that this new LMS could be very promising for odometry applications in typical road environments from night to day time, showing good precision in comparison with standard techniques when estimating the robot's position and orientation by using only two of these sensors and an EKF based on a simplified identification and calibration of the Ackermann model. However, further experiments along more complex and longer paths in several road conditions should be performed on a vehicle in order to fully evaluate the real potential of these sensors for odometry applications. In any case, as already said for automatic parking applications, the information provided by other sensors (wheel encoders, IMUs, GPSs, etc) should be combined together in order to achieve greater robustness, as required for the autonomous vehicles of the future.

Chapter 5

Conclusions & Perspectives

5.1 Accomplished work

In this thesis, we wanted to propose new bio-inspired technological solutions that could be used in advanced driver assistance systems (ADASs), and in particular on the autonomous vehicles of the future, for which redundancy and complementarity of the information perceived will be the main requirement for satisfying both autonomy and safety constraints.

As discussed in the first chapter of this manuscript, existing ADAS systems already offer a number of degrees of assistance to the driver, but, in their current form, are not yet capable of providing self-driving experiences that are complete and cost-competitive. Some prototypes of the first autonomous vehicles have already been presented, but such solutions are still too complex and expensive, and often require too large computational resources to be onboard standard cars.

We believe that effective technological solutions could be found by looking at the nature, which often suggests smart but "simple" solutions allowing animals to process only the information needed for well-defined tasks. The local motion sensors (LMSs), designed and developed at our laboratory using a very few pixels and computational resources by taking inspiration from the fly's visual system, definitely represent a good example of this approach.

Although the LMSs proposed so far have proved their effectiveness in measuring 1-D optic flow in certain situations/environments and have successfully been used on mobile and aerial robots for performing obstacle avoidance and speed control, they still required some improvements to correctly operate on real vehicles in road environments. The main improvements required were related to (i) their robustness with respect to the lighting conditions, visual textures and vibrations, and (ii) their ability to deliver measurements at a high, quasi-constant refresh rate with a constant, fixed resolution.

Two aspects were dealt therefore in this thesis: a technological aspect, involving the adaptation of LMSs to the road environments and the vehicle operating conditions, and an application aspect, involving the effective use of LMSs for robotic and automotive applications.

- In the first part, we showed that by using a more robust time-of-travel method for computing the optic flow based on cross-correlation of neighboring pixels' signals, which was inspired by the Hassenstein-Reichardt model, the LMSs could deliver high-frequency measurements when used on a vehicle driving along parked vehicles, even in presence of vibrations and few contrasts on the cars bodies. We also showed that, in situations where the LMSs delivered relatively accurate measurements, it was possible to compute 2-D point clouds at the same high frequency, in which the shapes of the parked vehicles could be recognized. Therefore, we might conclude that, in the case where several LMSs could be used all around the vehicle, the lines corresponding to the parked vehicles could be found in real time and the vehicle's ego-pose could be estimated in a simple but robust way during the entire parking maneuver, as showed in simulation. However, the response of these sensors was still not satisfactory in extreme lighting conditions (from shaded to lightened areas, with light reflected on vehicles, at night time, etc), mainly because of the signals delivered by the pixels in such conditions.

- In the second part, we presented a new auto-adaptive silicon retina featuring a novel pixel, called M^2APix, that reproduces the auto-adaptive responses recorded in the retina of most animals and even humans. Because testing auto-adaptive pixels and comparing them with other pixels responses appeared to us not an easy and well-defined operation, we also proposed a new standard method for accurately characterizing pixels' responses to luminous changes of up to ± 3 decades in a 7-decade range, by using a single light source, called Lighting Box. The novel M^2APix pixel showed very promising results during these characterization tests, in particular showing no deviation of its response in presence of sudden, very large changes in light, which is a common issue of VLSI chips. However, further tests in real operating conditions had to be performed to prove their true effectiveness for optic flow measurements.

- In the third and last part, a new robust optic flow sensor based on the new auto-adaptive chip and the new correlation-based method was therefore constructed, after improving the OF algorithm for providing a constant OF resolution and for successful implementation on embedded electronics. This new LMS was successfully tested on a low-cost car-like robot, called BioCarBot, for estimating its velocity and steering angle, and therefore estimating its trajectory (odometry), in various operating conditions (light levels, textures, vibrations, etc) similar to those occurring on road vehicles. A reliable and accurate estimation of the vehicle's velocity is, in fact, crucial as it is used by most of ADAS systems: the effectiveness of ABS and ESC systems, for instance, strongly depends on this estimation, but also that of any system controlling the vehicle's trajectory based on odometry (e.g. existing IPA systems). With the experiments carried out on the BioCarBot robot we showed that the LMS developed here (i) was able to cope with the main issues encountered with previous LMSs, improving their response outdoors in unstructured environments, and (ii) might provide a very promising solution for accurate velocity estimation and therefore odometry-based applications.

The new visual sensors presented in this thesis are bio-inspired in the following aspects: (i) their novel pixels reproduce the auto-adaptive response present in most animals' retina [36, 42, 47], and (ii) the new method for computing the optic flow combines the properties of time-of-travel algorithms [60, 62] with those of correlation-based algorithms based on the Hasseinstein-Reichardt model [76, 123], both inspired by biological findings obtained on insects. In addition, although the car-like robot, and therefore the model presented in the last part was not bio-inspired, the visual odometry method used was also inspired by the biological findings obtained on insects, e.g. on bees and hunts [93, 146], which show that ventral optic flow is mostly used by these animals to estimate their trajectory.

5.2 Further improvements

Although some important improvements have been successfully applied on the optic flow sensors presented here, they still present some limitations mainly due to the higher computational load than previous solutions. In fact, as already mentioned, the new correlation-based method for computing the optic flow could provide improved performances with respect to the threshold-based method used previously, but this resulted in an increased computational cost and therefore larger resources required (e.g. from a micro-controller to a micro-processor). A full quantitative comparison between the 2 methods should be performed to be able to determine which method it is worth to use, depending on the application, the hardware and the cost requirements.

In addition, although this new method allow to obtain a high refresh rate, a wide OF range and high and constant OF resolution, up to now it was not possible to obtain all these advantages at the same time as it would result in a saturation of the CPU load. As discussed at the end of chapter 4, the performances of the OF sensor could be further improved by optimizing and implementing the algorithm, for instance, on a DSP or FPGA, allowing to obtain a very good compromise between high frequency, wide OF range and high OF resolution.

Although the method presented here was developed to measure optic flow only in one direction, since it has been used in applications (odometry) where the direction of motion was known, it could be improved to measure the optic flow in the 2 directions by applying relatively simple modifications (see Appendix C). The same computations, in fact, could be performed after switching the delayed and non-delayed signals (motion in the opposite direction), then the maximum cross-correlation coefficient could be taken along the 2 directions, determining therefore also the sign of the optic flow. However, such an improvement would result in a even higher computational load, and some tests should be performed to verify its reliability.

As mentioned at the end of chapter 3, a few improvements could be also carried out on a new version of the M^2APix pixel in order to increase its effective voltage range up to the conversion range of the ADC. In this way, the contrast resolution would be increased nearly two-fold, resulting to a minimum detectable contrast of 1% in the light-adapted condition and 3% in the light-adapting condition, and thus

optic flow measurements would be possible even in the case of very little textured surfaces.

All in all, the LMSs developed up to present have some limitations with respect to other visual sensors, which are intrinsic to either their function or the technology used: they need textured surfaces and relative motion, they deliver motion information only in one dimension and their precision and distance range depend on the lens characteristics, namely the interreceptor angle $\Delta\varphi$, which can not be adjustable as it can be done, for instance, in standard cameras. The two former are limits intrinsic to the function of these sensors, i.e. visual motion and contrasts are necessary conditions for delivering measurements, and therefore can not be avoided. Conversely, the two latter are, as of now, technological and methodological limits, which could be probably improved respectively by (i) using 2-D pixels arrays and adapting the method for computing the optic flow in 2-D, and (ii) using a more complex optical system to mechanically adjust the lens characteristics online similarly to standard cameras.

The methods proposed here for parking and odometry-based applications also have some limitations, mainly due to the hypothesis that have been made in other to allow to use only the information provided by the LMSs (e.g. the hypothesis that the vehicles do not move, or that there is no lateral skid). This approach is very convenient when applied to robots that have to accomplish relatively simple tasks in not very complex environments, and therefore do not need redundant information and high level of intelligence. However, as already said in the first chapter, the autonomous vehicles of the future, as well as any autonomous mobile systems that have to operate in complex, populated environments, will need redundant and complementary information coupled with high level of intelligence to accurately perceive the environment and be able to make predictions and decisions on their actions. Therefore, the methods proposed here should be further improved to integrate data provided by other embedded sensors and performed more accurate estimations and control strategies.

In addition, the estimation and control methods used in these applications are statistically robust but relatively simple, since they are based on simplified kinematic models mainly derived from the Ackermann steering model. This makes it possible to perform a real-time implementation onboard low-cost car-like robots, and eventually cars, using relatively simple and light electronics, while obtaining good performances in relatively simple environments. However, when the environments become more complex and the dynamic of the vehicle may be important, these methods should be adapted to more complex dynamic models, and a more accurate identification and calibration of such models should also be performed. In addition, as we saw that the standard deviation of the measurement noise depended on the measurements' amplitude, methods involving dynamically-changing parameters of the covariance matrices, such as the Chong-Kleman method [313], could be used to increase the robustness of the EKFs.

5.3 The Multi-M^{2}APix sensor

During this thesis, we also designed and developed a new multi-facet sensor based on M^{2}APix chips, called Multi-M^{2}APix, as shown in Fig. 5.1.

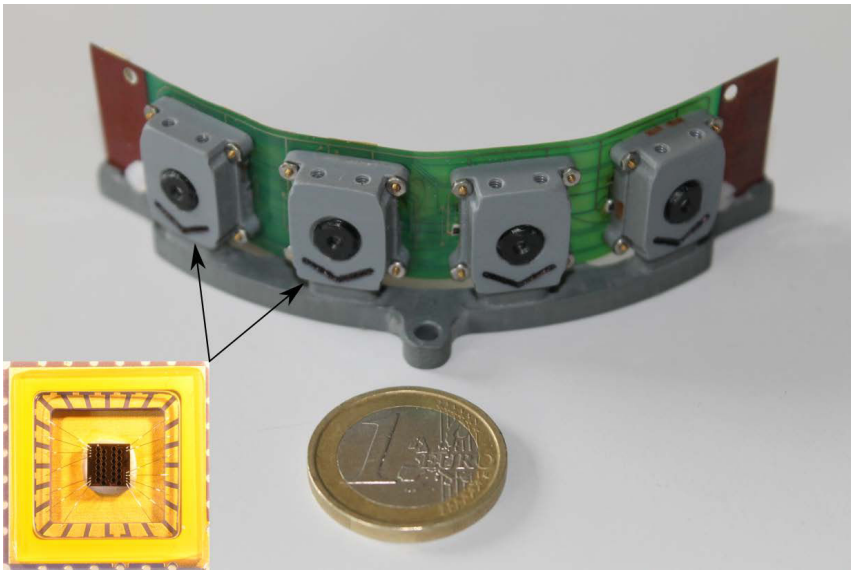

Figure 5.1: The Multi-M^{2}APix sensor composed of 4 M^{2}APix chips soldered on a flexible Printed Circuit Board (PCB), which can be flexed up to about 90^o between each chip/facet, allowing to eventually obtain a 360^o field of view (FOV). In this version, the lens used is the same to that used in the sensors presented in chapter 4, giving a $\sim 22^o$ FOV each and therefore a maximum contiguous FOV of about 90^o. However, this FOV can be customized by changing the lens, which can be easily unscrewed thanks to the 3-D printed optical casing mounted onto the chip.

The Multi-M^{2}APix sensor is composed of 4 M^{2}APix chips soldered on a 100x20mm flexible Printed Circuit Board (PCB), which can be flexed up to about 90^o between each chip/facet, allowing to eventually obtain a 360^o field of view (FOV). In the version developed here (Fig. 5.1), we used the same lens as that used in the sensors presented in chapter 4, giving a $\sim 22^o$ FOV each and therefore a maximum contiguous FOV of about 90^o. However, this FOV can be customized by changing the lens, which can be easily unscrewed thanks to the 3-D printed optical casing mounted onto the chips. This possibility to screw and unscrew the lens support also allow to adjust the distance between the lens and the pixels in order to obtain different spatial filtering, namely different values of $\Delta\varphi, \Delta\rho$, for each facet, reproducing similar characteristics to those of the insects' compound eyes but in a simpler and cheaper solution than that of artificial compound eyes (although much less compact).

Some details on the hardware implementation of this sensor are given in Appendix D.

This sensor was initially developed in order to evaluate the responses of multiple LMSs to compute 2-D point clouds all around the vehicle (or the car-like robot) during parking maneuvers. Then, in the case of good responses, the next step was to implement all the steps of the method proposed in chapter 2, and eventually perform closed-loop parking maneuvers based on the optic flow measurements delivered. Unfortunately, the development of this sensor encountered some difficulties and therefore took a too long time to be tested and used for these purposes.

5.4 Other applications and perspectives

Other applications that are not covered yet by existing ADAS systems have been also considered as interesting for use of multiple LMSs, such as automated driving close to walls, flank-guards or sidewalks, as well as in relatively narrow tunnels which can be found, for instance, in underground garages (Fig. 5.2). A robust estimation of the robot/vehicle's position and orientation with respect to the side walls (or flank-guards, sidewalks, etc) could be obtained, for instance, by adapting the solution proposed in chapter 2 for estimating the position and orientation of the lines found in the 2-D point clouds computed from the OF measurements. As the RANSAC method might be too computationally costly for implementation on low-cost embedded electronics, other solutions might be applied using EKFs based on the model given in equation (2.5) but, this time, using directly the OF measurements from neighboring pixels or the estimation of the line orientation obtained, for instance, by using the least squares method presented in [242].

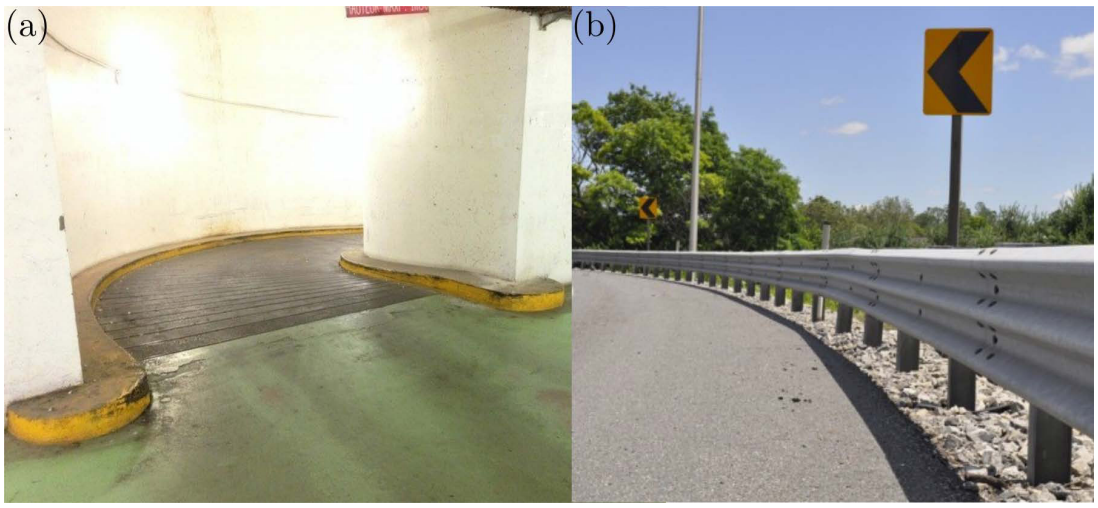

Figure 5.2: Examples of situations that are not covered yet by existing ADAS systems, in which LMSs could also be helpful for automated driving: (a) narrow tunnel in an underground garage; (b) a flank-guard at the side of a road.

Another application which could be considered is the detection of dangerous movements around the vehicle, in particular those produced by pedestrian, bikes or scooters, which are often not detected by the sensors used on ADAS available on the market due to their limits (see section 1.3). Such dangerous movements could be detected, for instance, by verifying the consistency in the OF measurements of pairs of neighboring pixels (or neighboring LMSs), or yet the consistency between the OF measurements delivered by the LMSs and some information (distances, velocities, etc) provided by other sensors or by some "a priori" knowledge on the environment.

Perspectives of how LMSs could equip standard cars were finally discussed in this thesis. Thanks to their simple structure and hardware requirements, we can imagine that these sensors could be integrated in any part of the car comprising transparent plastic elements, such as the headlights. In fact, these

sensors do not require a high resolution/precision of their optical lens, which is defocalized as opposed to standard cameras. Thus, we might propose that their lens could be directly produced by such transparent elements by locally deforming the plastic material to obtain a specific defocalization, and therefore a specific spatial filtering. Tiny PCBs including a few pixels would then be placed underneath this local plastic deformations (Fig. 5.3). In this way, the integration of these sensors would be achieved in a very cost-competitive way, giving local optic flow measurements all around the vehicle, which could therefore be used for some of the applications proposed all along this thesis.

These technical solutions have been described in details in 2 patents that have been recently filed (see the "Publications" page at the beginning of this manuscript).

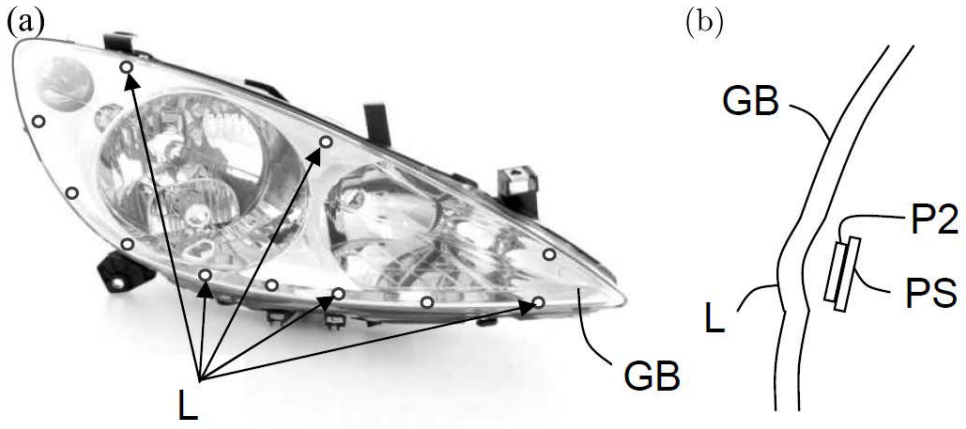

Figure 5.3: Example of integration of the LMSs on a car's headlight. (a) Several LMSs could be placed at different positions of the headlight and looking at different directions. (b) The lens (L) of these LMSs could produced by local deformations of the headlight's plastic (GB), and tiny PCBs (PS) comprising a few pixels (P2) could be then placed underneath these lens to obtain a specific defocalization, and therefore a specific spatial filtering.

All in all, the bio-inspired visual sensors presented in this thesis are complementary to the sensors already used on cars in the following aspects:

- they provide a higher dynamic range and higher frequency processing than standard cameras at lower cost than high-dynamic-range and high-frequency cameras, however they give only motion information in 1-D;

- they provide a wider field of view and longer distance range than ultrasonic sensors, however they give motion information instead of distance information;

- they provide a higher sensitivity to tangential movements (i.e. perpendicular to the sensor' axis) than radial movements (i.e. along the sensor's axis) as opposed to radars;

- they are passive sensors without any mechanical parts as opposed to lidars.

Therefore, these sensors might be used in autonomous vehicles for high-frequency motion detection around the vehicle, and in particular in the lateral hemispheres (e.g. in automated driving in parking garage), whereas cameras are still used for scene interpretation and object recognition at relatively low frequency in most applications, sonars are used for stationary and quasi-stationary object detection relatively close to the vehicle (e.g. in parking maneuvers) and radars are used for object detection relatively far from the vehicle in the front and rear hemispheres (e.g. in highway automated driving).

In addition, as shown in the last part of this thesis, the LMSs could also improved the odometry performances obtained when using GPSs, IMUs and wheel encoders.

Appendix A

Additional information on Chapter 2

Details on point and line motion modeling

Here, we describe the main steps which led to equations (2.1) and (2.5) in section 2.2.4.

Let us assume that the vector $\boldsymbol{X}_P$ represents the position of the point P with respect to the inertial frame $< I >$ (see Fig. 2.3), then we have:

$$ {}^I\boldsymbol{X}_P = {}^I\boldsymbol{X} + {}^I\boldsymbol{x}_P = {}^I\boldsymbol{X} + {}^IR_B({}^B\boldsymbol{x}_i + {}^B\boldsymbol{x}_s), \tag{5.1} $$

where $\boldsymbol{X}$ is the position of the of the body frame $< B >$ with respect to $< I >$, the upper-left subscripts indicate that the vectors are projected onto $< I >$ or $< B >$, and ${}^IR_B$ is the rotation matrix projecting the vectors from $< B >$ to $< I >$ (and viceversa for ${}^BR_I = {}^IR_B^T$).

By derivating equation (5.1) temporally with respect to $< I >$ and projecting the resulting vectors onto $< B >$, we obtain:

$$ \boldsymbol{V}_P = {}^BR_I \frac{d}{dt}\left({}^I\boldsymbol{X}_P\right) = {}^BR_I \frac{d}{dt}\left({}^I\boldsymbol{X}\right) + {}^BR_I \frac{d}{dt}\left({}^IR_B({}^B\boldsymbol{x}_s + {}^B\boldsymbol{x}_i)\right), \tag{5.2} $$

where $\boldsymbol{V}_P$ represents therefore the velocity of P with respect to $< I >$ projected onto $< B >$.

We note that the first term in the right-hand part of equation (5.2) is nothing but the longitudinal velocity of the vehicle $\boldsymbol{V}_f = [V_f\, 0\, 0]^T$, whereas the second term can be decomposed as follows:

$$ \begin{aligned} {}^BR_I \frac{d}{dt}\left({}^IR_B({}^B\boldsymbol{x}_s + {}^B\boldsymbol{x}_i)\right) &= {}^BR_I \left(\frac{d}{dt}\left({}^IR_B\right){}^IR_B^T {}^IR_B({}^B\boldsymbol{x}_s + {}^B\boldsymbol{x}_i) + {}^IR_B \frac{d}{dt}\left({}^B\boldsymbol{x}_s + {}^B\boldsymbol{x}_i\right) \right) \\ &= {}^BR_I \left(\hat{\Omega}\, {}^IR_B({}^B\boldsymbol{x}_s + {}^B\boldsymbol{x}_i) + {}^IR_B\, {}^B\dot{\boldsymbol{x}}_i \right) = \hat{\Omega}({}^B\boldsymbol{x}_s + {}^B\boldsymbol{x}_i) + {}^B\dot{\boldsymbol{x}}_i, \end{aligned} \tag{5.3} $$

where $\hat{\Omega} = \frac{d}{dt}\left({}^IR_B\right){}^IR_B^T$ is the skew-symmetric matrix of the vector $\boldsymbol{\Omega} = [0\, 0\, \dot{\theta}]^T$ corresponding to the angular velocity of $< B >$ with respect to $< I >$, $\frac{d}{dt}\left({}^B\boldsymbol{x}_s\right) = 0$ since ${}^B\boldsymbol{x}_s$ does not change, and where we used the properties ${}^IR_B^T\, {}^IR_B = I$ and ${}^BR_I \hat{\Omega}\, {}^IR_B = {}^BR_I\, {}^IR_B \hat{\Omega}$, with I being the identity matrix.

By combining equations (5.2) and (5.3) and rearranging the terms on the two sides, we obtain:

$$ {}^B\dot{\boldsymbol{x}}_i = \boldsymbol{V}_P - \boldsymbol{V}_f - \hat{\Omega}({}^B\boldsymbol{x}_s + {}^B\boldsymbol{x}_i), \tag{5.4} $$

which corresponds to equation (2.1) if we consider $\dot{\theta} = \frac{\tan\phi}{L} V_f$, according to the Ackermann model.

To obtain the first 2 equations in (2.5), let us consider that, for a sake of simplicity, the X-axis of $< I >$ is located along the line l. Then, we can write:

$$ {}^{B}\boldsymbol{x}_l = {}^{B}R_I \, {}^{I}\boldsymbol{x}_l = {}^{B}R_I \begin{bmatrix} 0 \\ -Y \end{bmatrix}, \tag{5.5} $$

where Y is the position of $< B >$ along the Y-axis (i.e. the second component of $\boldsymbol{X}$).

By derivating equation (5.5) temporally with respect to $< I >$, we obtain:

$$ {}^{B}\dot{\boldsymbol{x}}_l = {}^{B}R_I \frac{d}{dt}\left({}^{I}\boldsymbol{x}_l\right) + \frac{d}{dt}\left({}^{B}R_I{}^{I}\right) {}^{B}R_I{}^{I\,T} \, {}^{B}R_I{}^{I} \, {}^{I}\boldsymbol{x}_l = {}^{B}R_I{}^{I}\dot{\boldsymbol{x}}_l - \hat{\Omega}\,{}^{B}\boldsymbol{x}_l, \tag{5.6} $$

which corresponds to equation (2.5) if we consider that $\dot{Y} = \sin\theta V_f$ and $\dot{\theta} = \frac{\tan\phi}{L} V_f$, according to the Ackermann model.

In order to apply an extended Kalman filter (EKF) to the system described by equation (2.5), taking $\boldsymbol{x}_l$ as output, we need to verify its observability. The first 2 components of the state $\boldsymbol{\xi} = [x_l \, y_l \, \theta]^T$ are obviously observable since they correspond directly to the output $\boldsymbol{x}_l$, however we do not know whether the third variable is observable or not. For this purpose, we need to compute the rank of the extended observability matrix H_o, obtained by taking the extended output vector composed of the output itself and its time derivatives up to the $n-1$-th order (n being the dimension of state), and then computing the partial derivatives of this vector with respect to state variables. If the rank of this matrix is equal to the dimension of the system, i.e. $rank(H_o) = 3$, then the entire system, and therefore its state variables, is observable.

By derivating the output $\boldsymbol{x}_l$ with respect to time, we obtain the following observability matrix:

$$ H_o = \frac{\partial}{\partial \boldsymbol{\xi}} \begin{bmatrix} \boldsymbol{x}_l \\ \dot{\boldsymbol{x}}_l \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{\tan\phi}{L} V_f & -\sin(2\theta) V_f \\ -\frac{\tan\phi}{L} V_f & 0 & -\cos(2\theta) V_f \end{bmatrix}, \tag{5.7} $$

which is full rank since there are no values of the state and the control parameters (V_f, ϕ) that make the determinants of the 4 minors equal to zero at the same time.

Details on the experimental results

The vehicle's trajectory (${}^{I}\boldsymbol{X}$) shown in Fig. 2.8 was computed by integrating the equations of the Ackermann model [264] using the vehicle's velocity and steering angle values acquired via the vehicle's CAN bus as actual inputs. The 2-D point positions in the inertial frame $< I >$ (${}^{I}\boldsymbol{X}_P$) were then computed by using equation (5.1) obtaining the point clouds shown in Fig. 2.8.

Appendix B

Additional information on Chapter 3

Details on the chip

Here, we provide some details on the chip presented in chapter 3. The inputs and outputs of the silicon retina as well as the circuital scheme of the reference voltage regulator allowing the correct operation of the M^{2}APix pixel (see section 3.2.6) are shown in Fig. B.1.

A simplified description of the chip's inputs and outputs is given in table shown Fig. B.2. The reader is also suggested to look at Fig. 3.3 and 3.4 for a complete overview of the chip.

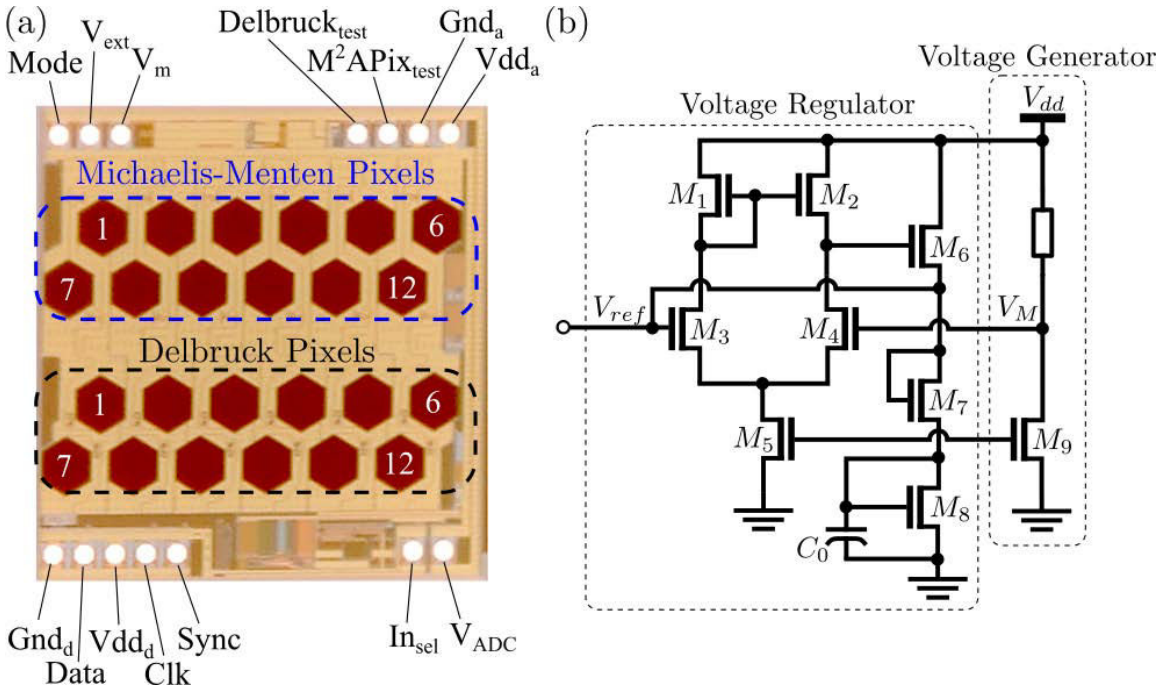

Figure B.1: (a) Magnified view of the silicon retina with its inputs and outputs. (b) Hardware implementation in VLSI of the reference voltage regulator allowing circuit operation over 7 decades of the photodiode current without causing saturation of the transistors (see Fig. 3.4).

Name	Type	Description
Mode	Digital Input	Value for internal (0) or external (1) normalizer current for the M²APix pixel
V_{ext}	Analog Input	Value corresponding to the external current for the M²APix pixel (if Mode = 1)
V_m	Analog Output	Value corresponding to the internal mean current for the M²APix pixel (a external capacitor C_m must be placed at this level)
$Delbruck_{test}$	Analog Output	Value corresponding to the 6th Delbruck pixel (for testing purposes)
M^2APix_{test}	Analog Output	Value corresponding to the 6th M²APix pixel (for testing purposes)
Gnd_a	Analog Input	Ground for the analog part (0 V)
Vdd_a	Analog Input	Voltage supply for the analog part (3.3 V)
Gnd_d	Analog Input	Ground for the digital part (0 V)
Vdd_d	Analog Input	Voltage supply for the digital part (3.3 V)
Data	Digital Output	Serial output of the pixels' values
Clk	Digital Input	Clock for sampling and serial read-out (max 1 MHz)
Sync	Digital Input	Trigger for the serial read-out interface
In_{sel}	Digital Input	Value for selecting M²APix (0) or Delbruck (1) pixel
V_{ADC}	Analog Input	High reference value for the ADC (from 1.2 V to Vdd_a)

Figure B.2: Table describing the chip's inputs and outputs.

Figure B.3 shows the table and the graph giving the theoretical conversion values from the external voltage V_{ext} to the current I_{ext} which goes into the normalizer block of the M²APix pixel when the $Mode$ parameter is equal to 1.

Luminosity	Voltage V_{ext}	Conversion V_{ext} to I_{est}
1 Lux (16 pA)	258 mV	0,47 pA/mV
10 Lux	340 mV	4,7 pA/mV
100 Lux	424 mV	42,5 pA/mV
1 kLux	520 mV	0,36 nA/mV
10 kLux	640 mV	2,65 nA/mV
10^5 Lux (1,6µA)	850 mV	11,5 nA/mV

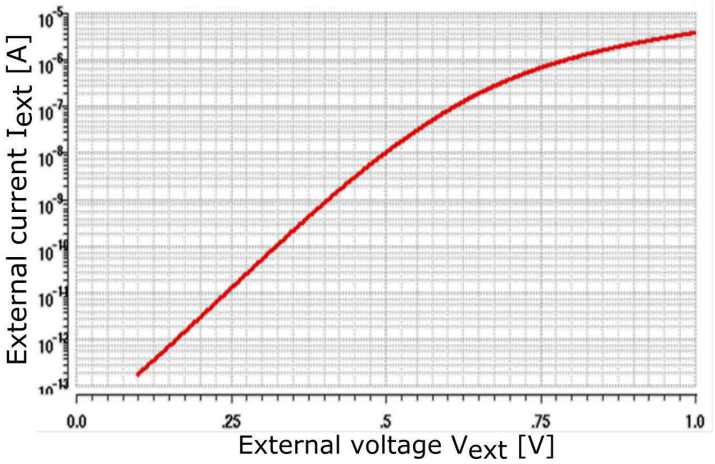

Figure B.3: Table and graph giving the theoretical conversion values from the external voltage V_{ext} to the current I_{ext}.

An internal state machine similar to that adopted in the CurvACE sensor [219, 242, 275] was implemented on chip for serial communication with an external device (Fig. B.4). When the chip is switched on or reset, a clean initialization of all the used blocks is performed by means of a power-on-reset circuit, therefore the first reading just after a reset should be not used by the external device. The sync and clock signals are initially pulled up using internal pulled up resistors, then four synchronized combinations of these signals impose the following states:

- **Wait mode**. After a power on or reset operation, the output data are all at the low level. This state is not changed while the sync signal stays at the high level. In this mode, the clock variation is not desirable to keep the power consumption at a lower level.

- **Full scan mode**. In order to set up this conversion mode, the sync signal has to go from 1 to 0 (falling edge) while the clock still at the high level. In this case, successive ADC conversions of the pixels' signals are performed and serial data are sent continuously to the external device at the clock frequency until the sync signal changes (from 0 to 1). In order to jump to the single scan mode, the sync signal has to change only at the low level of the clock and after an entire word (16 clock pulses) is sent. Otherwise the state machine reaches the stop mode.

- **Single scan mode**. In this mode, a counter disable signal is generated to take data only from the selected pixel. For example, if we need to digitize only the pixel 7, 6 words (6x16 clock pulses) have to be sent but only the result of the conversion of the 7th word will be sent on the serial bus (data output signal). The single scan mode can only be set up after at least one conversion in the full scan mode. This mode is mainly used for testing and characterization purposes where the converted data are generated, for instance, only for the first pixel.

- **Stop mode (or soft shutdown)**. The chip enters on this mode when one of these two conditions are satisfied: a sync transition from 0 to 1 occurs while the clock is at the high level if the machine is in full scan mode, or from 1 to 0 occurs while the clock is at the low level if the machine is in single scan mode. In this mode, all internal blocks can be powered down. If all the analog blocks are turned off, then the external device will have to wait for a while, namely the weak up time, before validating data from conversion. If the blocks stayed switched on, the state machine goes back to the wait mode.

The 16-bit words sent on the serial output data are composed as follows: 0 YYYY XXXXXXXXXX 0, where the 4 Y bits define the pixel number starting from the most significant bit, i.e. 1000 for pixel 1, 0100 for pixel 2, etc (0000 corresponds to V_m for the M^{2}APix pixel and to the bandgap voltage for the Delbruck pixel), and the 10 X bits represent the pixel value starting from the least significant bit.

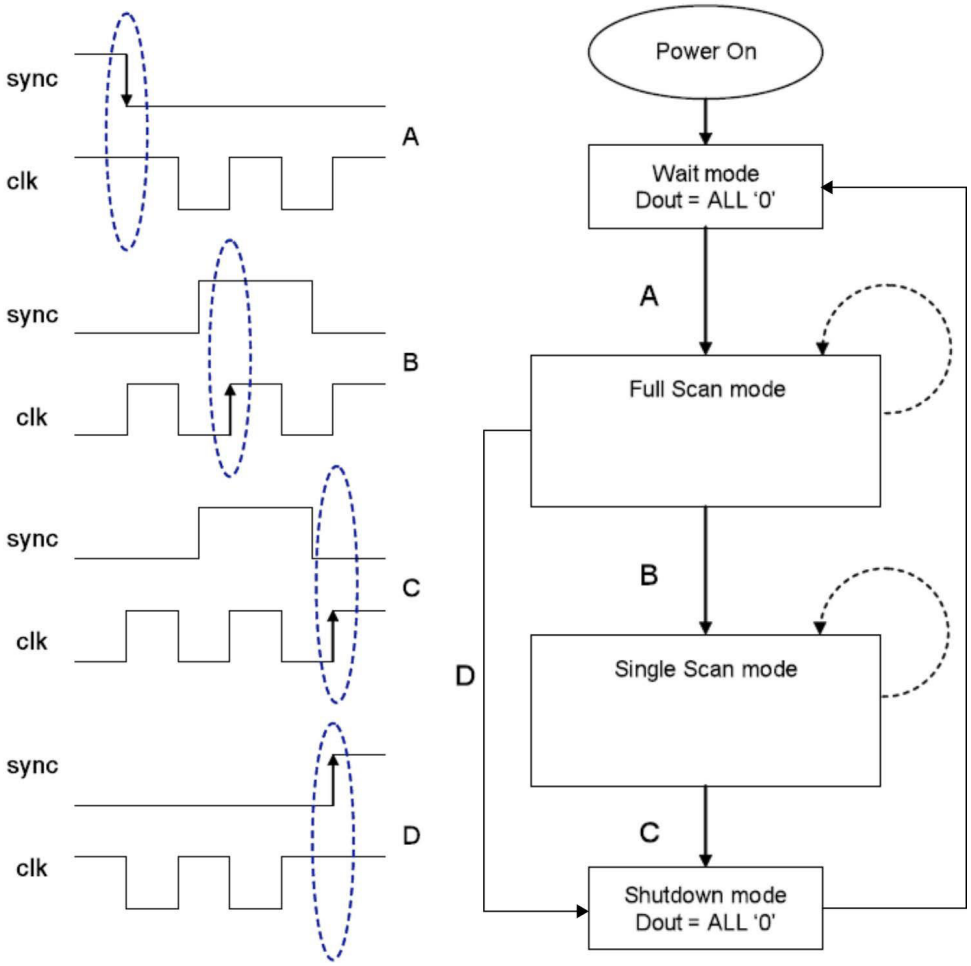

Figure B.4: Temporal and block diagram of the state machine implemented on the chip [275]. Courtesy of Center of Particle Physics of Marseille (CPPM).

Characterization of the Delbruck pixel

Here, we present the results obtained for the characterization of the Delbruck pixels present on our silicon retina, in addition to the M^{2}APix pixels (Fig. B.1). The same hardware setup and characterization procedure have been used as for the M^{2}APix pixels (see section 3.2.8). However, only one optical filter has been used for each average luminosity, therefore the s-shaped curves were characterized over ± 1.5 decades instead of ± 3 decades as for the M^{2}APix pixel (Fig. B.5).

This characterization showed that, apart for the pixel 6 which did not properly respond (yellow line with circle markers in Fig. B.5), the 12 pixels qualitatively reproduced the expected s-shaped curved as provided in [175] at relatively low luminosity, whereas they delivered unexpected and unrepeatable responses for large changes in light (> 1 decade, corresponding to contrasts higher than 50%) at relatively high luminosity (Fig. B.5). Such an unreliable response for large changes in light represents the main drawback of the Delbruck pixel and, on the contrary, a major advantage of the M^{2}APix pixel, as already discussed in section 3.2.10.

Another difference between the Delbruck and M^{2}APix pixel is the steady-state response: in the former it increases with the luminosity, resulting in a theoretically unbounded output and a non-optimal LSB resolution, since the transient operating range at each light level ($\sim 1\,V$) is smaller than the whole operating range ($\sim 1.8\,V$), whereas in the latter it is nearly constant independently to the luminosity, resulting in a limited-range output and therefore an improved LSB resolution.

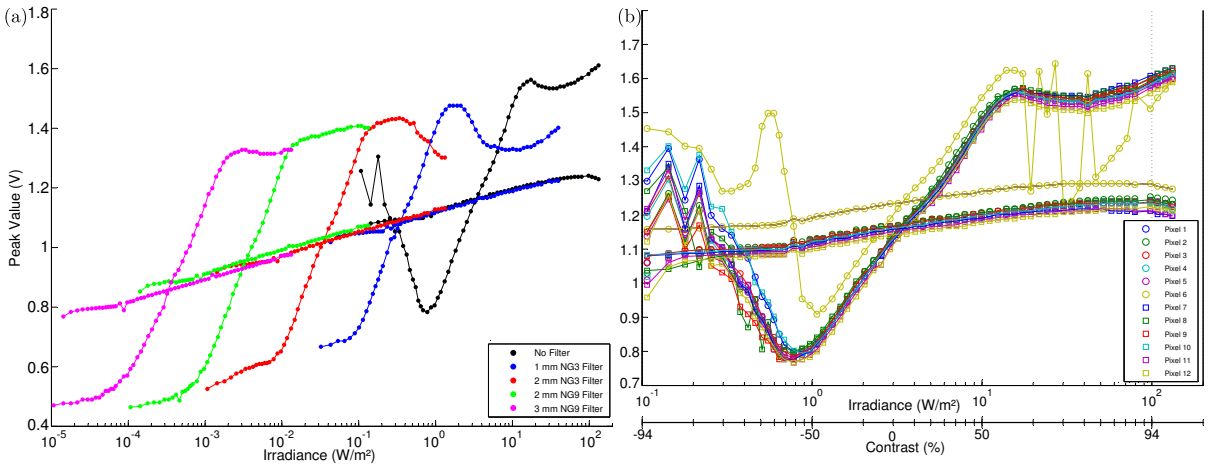

Figure B.5: S-shaped curves and steady-state responses of the pixels of Delbruck type present on our silicon retina. (a) Peak and steady-state values obtained for the third pixel at different average luminosities using four optical filters (different colors, black is without optical filter). (b) Peak and steady-state values obtained for the 12 pixels (different colors and markers) at one average luminosity (no optical filter).

Appendix C

Additional information on Chapter 4

Algorithm implementation of the new optic flow method

Here, we describe the algorithm used for the correlation-based optic flow method presented in chapter 4, which has been improved to reduce the computational cost and therefore be implemented on embedded electronics.

The Pearson cross-correlation coefficient between 2 sampled signals x, y at one sampling step can defined as follows:

$$\varrho = \frac{cov(x,y)}{\sigma_x \sigma_y} = \frac{\sum_i^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i^n (x_i - \bar{x})^2} \sqrt{\sum_i^n (y_i - \bar{y})^2}}, \tag{5.8}$$

where $\bar{x} = \frac{\sum_i^n x_i}{n}$ is the mean value of x (same for $\bar{y}$), the index i indicates the i-th sample and n is the number of samples. The sampling time Δt and the number of samples n are constant and fixed, and therefore x, y can be considered as FIFO buffers of fixed length n.

In our case, the ϱ coefficients are computed for N pairs of signals x, y, corresponding to the pixels' signals $V_{Ph_j}, V_{Ph_{j+1}}$ ($N = 10$ for the M^{2}APix sensor, see section 4.2.3), and each pair for m time lags, giving $N \times m$ computations of equation (5.8) and therefore $N \times m \times n$ subtractions and products for each signal. Such a computation can be too costly for embedded electronics if we want to have robust measurements (high value of n) in a wide OF range with high OF resolution (high value of m).

By looking at equation (5.8), we can note that, as a new sample enters in the FIFO buffers, the sums $\sum_i^n (x_i - \bar{x})$ and $\sum_i^n (y_i - \bar{y})$ (and their squared counterparts) computed from one sampling step to the next one have $n - 2$ terms unchanged, i.e. the terms corresponding to $i = 2, \ldots, n-1$. Therefore, the computation of the ϱ coefficients at each sampling step can be improved by updating these sums by subtracting the n-th term (corresponding to the oldest sample in the FIFO) and adding the first term (corresponding to the new sample entering in the FIFO), instead of recomputing them from zero. In this way, we will have about $N \times m \times 2$ subtractions and products instead of $N \times m \times n$.

Therefore, the improved pseudo-algorithm applied to each pair of pixels' signals V_{Ph_1}, V_{Ph_2} sampled, band-pass filtered and buffered in the FIFO buffers Ph_1, Ph_2, can be described as follows:

Algorithm 1 Improved cross-correlation method for computing the optic flow

function OF(Ph_1, Ph_2) ▷ Ph_1, Ph_2 are FIFO buffers
  $Ph'_2 \leftarrow Ph_1[n+1]$ ▷ Take the oldest sample of Ph_2
  $Ph''_2 \leftarrow Ph_2[1]$ ▷ Take the newest sample of Ph_2
  $Ph'_{2_{mean}} \leftarrow \frac{Ph_{2_{sum}}}{n}$ ▷ $Ph'_{2_{mean}}$ is the previous average of Ph_2 ($Ph_{2_{sum}}$ is initialized at 0)
  $Ph_{2_{sum}} \leftarrow Ph_{2_{sum}} - Ph''_2 + Ph'_2$ ▷ Update the sum of the elements of Ph_2
  $Ph''_{2_{mean}} \leftarrow \frac{Ph_{1_{sum}}}{n}$ ▷ $Ph''_{2_{mean}}$ is the updated average of Ph_2
  $C'_2 \leftarrow Ph'_2 - Ph'_{2_{mean}}$
  $C''_2 \leftarrow Ph''_2 - Ph''_{2_{mean}}$
  $C_2 = C_2 - C'^2_2 + C''^2_2$ ▷ C_2 is initialized at 0
  for $i \leftarrow 1, \ldots, m$ **do**
    $Ph'_1 \leftarrow interp(Ph_1, \tau_i, n+1)$ ▷ Take the oldest subsamples of Ph_1 starting from $t = \tau_i$
    $Ph''_1 \leftarrow interp(Ph_1, \tau_i, 1)$ ▷ Take the newest subsamples of Ph_1 starting from $t = \tau_i$
    $Ph'_{1_{mean}} \leftarrow \frac{Ph_{1_{sum}}}{n}$ ▷ $Ph'_{1_{mean}}$ is the previous average of Ph_1 ($Ph_{1_{sum}}$ is initialized at 0)
    $Ph_{1_{sum}} \leftarrow Ph_{1_{sum}} - Ph''_1 + Ph'_1$ ▷ Update the sum of the elements of Ph_1
    $Ph''_{1_{mean}} \leftarrow \frac{Ph_{1_{sum}}}{n}$ ▷ $Ph''_{1_{mean}}$ is the updated average of Ph_1
    $C'_1 \leftarrow Ph'_1 - Ph'_{1_{mean}}$
    $C''_1 \leftarrow Ph''_1 - Ph''_{1_{mean}}$
    $C_1 = C_1 - C'^2_1 + C''^2_1$ ▷ C_1 is initialized at 0
    $C_{12} = C_{12} - C'_1 C'_2 + C''_1 C''_2$ ▷ C_{12} is initialized at 0
    $\varrho_i \leftarrow \frac{C_{12}}{\sqrt{C_1 C_2}}$ ▷ Compute the Pearson cross-correlation coefficient
  end for
  $k \leftarrow \arg\max_i \varrho_i$ ▷ Find the maximum correlation
  if $\varrho_k > \varrho_{thr}$ **then**
    $\omega \leftarrow \pm\frac{\Delta\varphi}{\tau_k}$ ▷ The sign depends on the sensor's axis orientation
  else
    $\omega \leftarrow NaN$
  end if
  return ω
end function

The algorithm presented above could be extended to measure the optic flow in 2 directions by applying a nested *for* loop ($j \leftarrow 1, 2$), in which the same computations (from line 2 to 21) would be performed first by delaying Ph_2, as in the case shown here, and then by delaying Ph_1. The maximum cross-correlation coefficient would be then taken along the 2 indexes ($i = 1, \ldots, m$ and $j = 1, 2$), and the sign of ω would be therefore determined by the index j, i.e. $k, h \leftarrow \arg\max_{i,j} \varrho_{i,j}$ and $\omega \leftarrow \pm\frac{\Delta\varphi}{\tau_k}$ depending on whether $h = 1$ or 2.

Details on the BioCarBot robot

Figure C.1 shows a functional diagram of the electronics used on the BioCarBot robot.

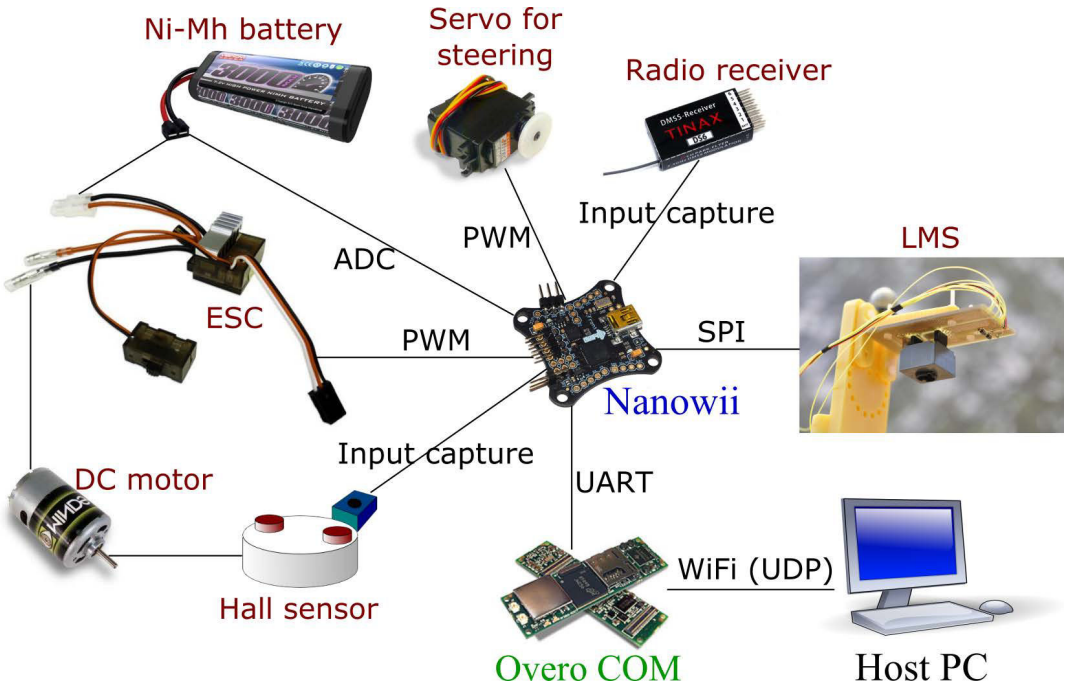

Figure C.1: Functional diagram of the electronic hardware and connections implemented on the BioCarBot robot.

As mentioned in sections 4.2.4 and 4.2.5, the main electronics included one Nanowii board (*Flyduino*) featuring a ATmega32u4 16-MHz CPU microcontroller (*Atmel*) and a MPU-6050 inertial measurement unit (IMU) comprising a 3-axis gyroscope and a 3-axis accelerometer (*InvenSense*) and one Overo IronSTORM computer-on-module (COM) (*Gumstix*) featuring a 1-GHz CPU DM3730 processor (*Texas Instruments*) comprising an ARM Cortex-A8 architecture and a C64x digital signal processor (DSP). The M^{2}APix output signals were sampled at a frequency of 333 Hz ($\Delta t = 3\,ms$) by the on-chip ADC (see [261] for details), acquired by the Nanowii board via SPI communication and transmitted to the COM via serial communication (UART). The Nanowii board also transmitted to the COM other values, such as the IMU values, the motor speed values measured by a hall sensor attached to the motor shaft, and the battery voltage level measured via an ADC, while receiving speed and steering control values. The OF algorithm

presented above as well as the estimation and control scheme shown in Fig. 4.5 were run at the same rate on the COM. The ground host PC sends the control set points to the robot's COM and receives data from the robot's COM and the Vicon system through a Wi-Fi connection via UDP communication.

In addition, the Nanowii board was interrupted every $0.02\,s$ (50 Hz) by a radio receiver, which received 3 commands from a remote controller: manual or automatic mode selection, and speed and steering control values. Then, the Nanowii board transmitted at the same frequency the desired speed and steering values to the electronic speed controller (ESC) and the servo via pulse width modulated (PWM) signals. These values were those received by the radio receiver if the mode selected was manual, otherwise were those received by the COM via serial communication (UART).

Supplementary Data of the Article

	$\Delta\omega^*$	$\bar{f}_\omega$	$\frac{\sigma^l_{\omega^l_m}}{\bar{\omega}^l_m}$, $\frac{\sigma^r_{\omega^r_m}}{\bar{\omega}^r_m}$ $(\bar{\omega}^l_m, \bar{\omega}^r_m)$	$\frac{\sigma_V}{\bar{V}}$, $\frac{\sigma_\phi}{\lvert\bar{\phi}\rvert}$ $(\bar{V}, \lvert\bar{\phi}\rvert)$	$\frac{\lvert\lvert\mathbf{X}-\hat{\mathbf{X}}\rvert\rvert_1}{d}$, $\frac{\lvert\theta-\hat{\theta}\rvert}{\lvert\theta\rvert}$ $(d, \lvert\theta\rvert)$	$\frac{\lvert\lvert\mathbf{X}-\hat{\mathbf{X}}_{IMU}\rvert\rvert_1}{d}$, $\frac{\lvert\theta-\hat{\theta}_{IMU}\rvert}{\lvert\theta\rvert}$ $(d, \lvert\theta\rvert)$
High-dynamic-range lighting conditions (Fig. 4.6)	$0.3\ \frac{rad}{s}$	$327\,Hz$	3.3%, 3.2% $(4.5,\ 5\ \frac{rad}{s})$	3%, 12% $(0.8\ \frac{m}{s},\ 0.15\,rad)$	2%, 1.2% $(45\,m,\ 8.1\pi\,rad)$	8.4%, 10.1% $(44\,m,\ 8.1\pi\,rad)$
High OF resolution (Fig. 4.7(A))	$0.05\ \frac{rad}{s}$	$333\,Hz$	1.2%, 1.4% $(7.5,\ 9\ \frac{rad}{s})$	0.9%, 6% $(1.5\ \frac{m}{s},\ 0.15\,rad)$	2.5%, 1.5% $(18\,m,\ 4.1\pi\,rad)$	2%, 1.5% $(22\,m,\ 4.1\pi\,rad)$
Wide OF range (Fig. 4.7(B))	$0.5\ \frac{rad}{s}$	$332\,Hz$	3.6%, 3.4% $(7.5,\ 8.5\ \frac{rad}{s})$	2.4%, 12% $(0.8\ \frac{m}{s},\ 0.15\,rad)$	2.2%, 2.2% $(45\,m,\ 8.6\pi\,rad)$	4.3%, 4.6% $(43\,m,\ 8.6\pi\,rad)$
Eight-shaped path (Fig. 4.8(A))	$0.1\ \frac{rad}{s}$	$326\,Hz$	2.3%, 2.1% $(4,\ 4\ \frac{rad}{s})$	1.8%, 5.3% $(0.6\ \frac{m}{s},\ 0.3\,rad)$	2.3%, 1.8% $(31\,m,\ 6.1\pi\,rad)$	6.1%, 6% $(32\,m,\ 6.1\pi\,rad)$
Square-shaped path (Fig. 4.8(B))	$0.2\ \frac{rad}{s}$	$324\,Hz$	3.1%, 3.3% $(3.5,\ 3\ \frac{rad}{s})$	3.3%, 11% $(0.6\ \frac{m}{s},\ 0.2\,rad)$	0.7%, 0.3% $(71\,m,\ 16.6\pi\,rad)$	1.6%, 1.7% $(71\,m,\ 16.6\pi\,rad)$

Table 5.1: Quantitative data and statistics on the indoor tests.

	$\Delta\omega^*$	$\bar{f}_\omega$	$\frac{\sigma^l_m}{\bar{\omega}^l_m}$, $\frac{\sigma^r_m}{\bar{\omega}^r_m}$ $(\bar{\omega}^l_m, \bar{\omega}^r_m)$	$\frac{\sigma_V}{\bar{V}}$, $\frac{\sigma_\phi}{\lvert\bar{\phi}\rvert}$ $(\bar{V}, \lvert\bar{\phi}\rvert)$	$\frac{\lvert\lvert\mathbf{X}-\hat{\mathbf{X}}\rvert\rvert_2}{d}$, $\frac{\lvert\theta-\hat{\theta}\rvert}{\lvert\theta\rvert}$ $(d, \lvert\theta\rvert)$	$\frac{\lvert\lvert\mathbf{X}-\hat{\mathbf{X}}_{IMU}\rvert\rvert_2}{d}$, $\frac{\lvert\theta-\hat{\theta}_{IMU}\rvert}{\lvert\theta\rvert}$ $(d, \lvert\theta\rvert)$
Uneven terrains at day light (Fig. 4.9)	0.3 $\frac{rad}{s}$	301 Hz	3.6%, 3.2% (4.5, 5 $\frac{rad}{s}$)	2.6%, 11% (0.8 $\frac{m}{s}$, 0.2 rad)	2.6%, 2.2% (31 m, 8π rad)	4%, 3.3% (30 m, 8π rad)
Uneven terrains at night (Fig. 4.10)	0.3 $\frac{rad}{s}$	299 Hz	3.8%, 3.6% (4.5, 5 $\frac{rad}{s}$)	3.1%, 12% (0.8 $\frac{m}{s}$, 0.2 rad)	3%, 2.3% (32 m, 8π rad)	4.8%, 4.1% (30 m, 8π rad)
Various ground textures (Fig. 4.11)	0.2 $\frac{rad}{s}$	295 Hz	3.7%, 3% (3, 4 $\frac{rad}{s}$)	2.2%, 9% (0.8 $\frac{m}{s}$, 0.2 rad)	3%, 1.9% (31 m, 8π rad)	3.3%, 3.9% (34 m, 8π rad)
Eight-shaped path (Fig. 4.12)	0.1 $\frac{rad}{s}$	328 Hz	2.9%, 2.6% (3.5, 3.5 $\frac{rad}{s}$)	1.8%, 7.3% (0.6 $\frac{m}{s}$, 0.3 rad)	1.9%, 1.6% (28 m, 6π rad)	8.3%, 8.5% (27 m, 6π rad)

Table 5.2: Quantitative data and statistics on the outdoor tests.

Appendix D

Additional information on the Multi-M^{2}APix sensor

Here, we provide some details on the Multi-M^{2}APix sensor shown in Fig. D.1.

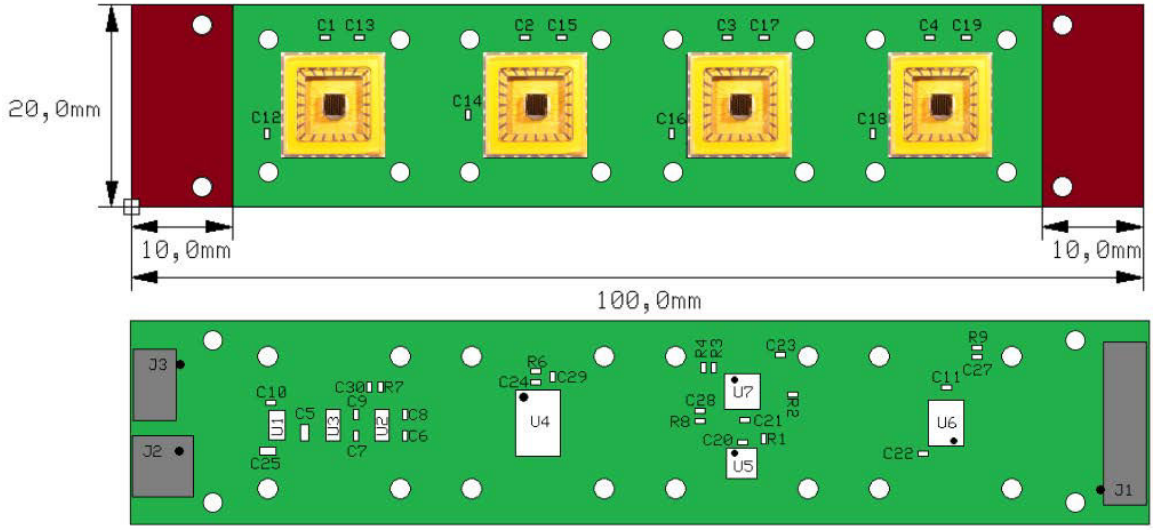

Figure D.1: Schematic view of the front and back side of the Multi-M^{2}APix sensor, composed of 4 M^{2}APix chips soldered on a 100x20mm flexible Printed Circuit Board (PCB) (green area), which can be flexed up to about 90^o between each chip/facet, allowing to eventually obtain a 360^o field of view (FOV). The 2 brown areas at the left and right-hand sides are rigid parts due to the connectors (grey blocks). The details on the connectors' pins are given in the table shown in Fig. D.2.

The 100x20mm flexible Printed Circuit Board (PCB) is composed of 4 M^{2}APix chips and some additional electronics that allow to adjust the parameters of each chip (see Appendix B) from an external device in a very portable way. The main electronic components included on the flex PCB are:

- one IO expander with 4 separate outputs driven by a single digital input via I^{2}C communication (PCA9536 by Texas Instruments), to set the parameter In_{sel} at 0 or 1 to select either the M^{2}APix or the Delbruck pixels;
- one digital potentiometers with 4 separate outputs driven by a single digital input via I^{2}C communication (AD5144A by Analog Devices), to adjust the voltage V_{ext} corresponding to the external

current I_{ext} used in the M^{2}APix pixels if $Mode = 1$[1] (as the operating voltage range is $[250, 850]\, mV$, a voltage regulator at $1\, V$ was also included on the PCB to enhance the potentiometer output resolution);

- 2 level translators (TXB0104 and TXB0108 by Texas Instruments), to allow a suitable interface with external devices operating at different voltage levels.

The pixels' data can be acquired via SPI communication by using either the 2 separate buses available (in this case, the chips are grouped by 2) or only a single bus by using a jumper on the connector. Due to the maximum frequency of integrated ADC (1 MHz), the data from the 4 chips (12x4 pixels) can be acquired at a maximum frequency of ~ 2.2 Hz and ~ 1.1 Hz, respectively in the 2 cases. Additionally, to enhance the pixels voltage conversion, a voltage regulator at $2.5\, V$ was also included on the flex PCB and connected to V_{ADC} input of each chip, setting the high reverence value of the integrated ADC.

The details on the connectors present on the flex PCB are given in the table in Fig. D.2, where SCL, SDA are the I^{2}C clock and data bus, respectively, for the IO expanders and the digital potentiometers; SCLK_X, MISO_X (X = 1,2) are the SPI clocks and data buses (MISO), respectively, for the 2 groups of chips (SCLK_1, MISO_1 can be connected to SCLK_2, MISO_2 for using a single bus); Sync_X (X = $1, \ldots, 4$) are the Sync signals for the read-out interface of each chip (corresponding to the "chip select" of the SPI protocol); Vcc is the voltage level of the external device needed for the level translators; Vm_X (X = $1, \ldots, 4$) are the voltage values corresponding to the internal mean current of M^{2}APix pixels of each chip (see Appendix B); Vin is the voltage supply which must be set at $3.3\, V$.

J1	***1***	***2***	***3***	***4***	***5***	***6***	***7***	***8***
Signals Ext	Gnd_Ext	Gnd_Ext	Mode_Ext	SCL_Ext	SDA_Ext	Sync1_Ext	Sync4_Ext	Sync3_Ext

J1 (cont)	***9***	***10***	***11***	***12***	***13***	***14***
Signals Ext	Sync2_Ext	MISO_1_Ext MISO_2_Ext	SCLK_1_Ext SCLK_2_Ext	SCLK_2_Ext	MISO_2_Ext	Vcc_Ext

J2	***1***	***2***	***3***
Power supply	Gnd	+ Vin	NC

J3	***1***	***2***	***3***	***4***	***5***
Signals Vm	Gnd	Vm_1	Vm_2	Vm_3	Vm_4

Figure D.2: Table of the connectors' pins of the Multi-M^{2}APix sensor shown in Fig. D.1.

[1] The $Mode$ parameter is set at 0 or 1 by the external device.

Bibliography

[1] I. Asimov, "Runaround," in *Astounding Science Fiction*, Street & Smith, 1942.

[2] N. Wiener, *Cybernetics or Control and Communication in the Animal and the Machine*. The Technology Press, 1948.

[3] N. J. Nilsson, "Shakey the robot," tech. rep., AI Center, SRI International, 1984.

[4] R. Siegwart, K. O. Arras, S. Bouabdallah, D. Burnier, G. Froidevaux, X. Greppin, B. Jensen, A. Lorotte, L. Mayor, M. Meisser, R. Philippsen, R. Piguet, G. Ramel, G. Terrien, and N. Tomatis, "Robox at expo.02: A large-scale installation of personal robots," *Robotics and Autonomous Systems*, vol. 42, no. 3-4, pp. 203–222, 2003. Socially Interactive Robots.

[5] J. Biswas and M. Veloso, "Localization and navigation of the cobots over long-term deployments," *Int. J. Rob. Res.*, vol. 32, no. 14, 2013.

[6] U.S. National Highway Traffic Safety Administration, "The long-term effect of abs in passenger cars and ltvs," 2009.

[7] U.S. National Highway Traffic Safety Administration, "Federal motor vehicle safety standards; electronic stability control systems," 2006.

[8] Peugeot, "Park assist." `http://www.feline208.net/dossiers/equipements/park-assist/`.

[9] Volkswagen, "Park steerig assistance." `http://www.volkswagenag.com/content/vwcorp/content/en/innovation/driver_assistance/parking_steering_assistance.html`.

[10] Toyota, "Intelligent park assist." `http://www.toyota.com.au/prius-v/features/audio-technology/intelligent-park-assist`.

[11] Nissan, "Around view monitor system." `http://www.nissan-global.com/EN/TECHNOLOGY/OVERVIEW/avm.html`.

[12] BMW, "Surround view system." `http://www.bmw.com/com/en/insights/technology/connecteddrive/2013/driver_assistance/intelligent_parking.html`.

[13] M. Montemerlo, J. Becker, S. Bhat, H. Dahlkamp, D. Dolgov, S. Ettinger, D. Haehnel, T. Hilden, G. Hoffmann, B. Huhnke, D. Johnston, S. Klumpp, D. Langer, A. Levandowski, J. Levinson, J. Marcil, D. Orenstein, J. Paefgen, I. Penny, A. Petrovskaya, M. Pflueger, G. Stanek, D. Stavens, A. Vogt, and S. Thrun, "Junior: The Stanford Entry in the Urban Challenge," *J. Field Robot.*, vol. 25, pp. 569–597, 2008.

[14] C. Urmson, J. Anhalt, D. Bagnell, C. Baker, R. Bittner, M. Clark, J. Dolan, D. Duggins, T. Galatali, C. Geyer, *et al.*, "Autonomous driving in urban environments: Boss and the urban challenge," *Journal of Field Robotics*, vol. 25, no. 8, pp. 425–466, 2008.

[15] H. Jung, Y. H. Cho, P. J. Yoon, and J. Kim, "Scanning laser radar-based target position designation for parking aid system," *IEEE Trans. Intell. Transp. Syst.*, vol. 9, pp. 406–424, 2008.

[16] U.S. National Highway Traffic Safety Administration, "Nhtsa announces final rule requiring rear visibility technology." `http://www.nhtsa.gov/About+NHTSA/Press+Releases/2014/NHTSA+Announces+Final+Rule+Requiring+Rear+Visibility+Technology`, 2014.

[17] Z. Tao, P. Bonnifait, V. Frémont, and J. Ibañez-Guzman, "Mapping and localization using gps, lane markings and proprioceptive sensors," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 406–412, 2013.

[18] V. Drevelle and P. Bonnifait, "Localization Confidence Domains via Set Inversion on Short-Term Trajectory," *IEEE Transactions on Robotics*, vol. 29, no. 5, pp. 1244–1256, 2013.

[19] D. Scaramuzza and R. Siegwart, "Appearance-guided monocular omnidirectional visual odometry for outdoor ground vehicles," *IEEE Transactions on Robotics*, vol. 24, no. 5, pp. 1015–1026, 2008.

[20] D. Scaramuzza, "Performance evaluation of 1-point-ransac visual odometry," *Journal of Field Robotics*, vol. 28, no. 5, pp. 792–811, 2011.

[21] G. Bleser and G. Hendeby, "Using optical flow as lightweight slam alternative," in *Mixed and Augmented Reality, 2009. ISMAR 2009. 8th IEEE International Symposium on*, pp. 175–176, 2009.

[22] P. Alcantarilla, J. Yebes, J. Almazan, and L. Bergasa, "On combining visual slam and dense scene flow to increase the robustness of localization and mapping in dynamic environments," in *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, pp. 1290–1297, 2012.

[23] Wikicars, "Driverless car." `http://wikicars.org/en/Driverless_car`.

[24] E. D. Dickmanns, R. Behringer, D. Dickmanns, T. Hildebrandt, M. Maurer, F. Thomanek, and J. Schiehlen, "The seeing passenger car 'vamors-p'," in *Intelligent Vehicles '94 Symposium, Proceedings of the*, pp. 68–73, 1994.

[25] E. D. Dickmanns, *Dynamic Vision for Perception and Control of Motion*. Springer-Verlag New York, Inc., 2007.

[26] E. D. Dickmanns, "Vehicles capable of dynamic vision: a new breed of technical beings?," *Artificial Intelligence*, vol. 103, no. 1-2, pp. 49–76, 1998.

[27] A. Broggi, M. Bertozzi, A. Fascioli, C. Guarino, L. Bianco, and A. Piazzi, "The argo autonomous vehicle's vision and control systems," *International Journal of Intelligent Control and Systems*, pp. 409–441, 1999.

[28] C. Urmson, J. Anhalt, M. Clark, T. Galatali, J. P. Gonzalez, J. Gowdy, A. Gutierrez, S. Harbaugh, M. Johnson-Roberson, H. Kato, *et al.*, "High speed navigation of unrehearsed terrain: Red team technology for grand challenge 2004," *Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, Tech. Rep. CMU-RI-04-37*, 2004.

[29] S. Thrun, M. Montemerlo, H. Dahlkamp, D. Stavens, A. Aron, J. Diebel, P. Fong, J. Gale, M. Halpenny, G. Hoffmann, K. Lau, C. Oakley, M. Palatucci, V. Pratt, P. Stang, S. Strohband, C. Dupont, L.-E. Jendrossek, C. Koelen, C. Markey, C. Rummel, J. van Niekerk, E. Jensen, P. Alessandrini, G. Bradski, B. Davies, S. Ettinger, A. Kaehler, A. Nefian, and P. Mahoney, "Stanley: The robot that won the darpa grand challenge," *Journal of Field Robotics*, vol. 23, no. 9, pp. 661–692, 2006.

[30] A. Broggi, P. Medici, P. Zani, A. Coati, and M. Panciroli, "Autonomous vehicles control in the vislab intercontinental autonomous challenge," *Annual Reviews in Control*, vol. 36, no. 1, pp. 161–171, 2012.

[31] Volkswagen, "Temporary auto pilot." `http://www.volkswagenag.com/content/vwcorp/content/en/innovation/driver_assistance/Temporary_Auto_Pilot.html`.

[32] T. M. Gasser, *Handbook of Intelligent Vehicles*, ch. Legal Issues of Driver Assistance Systems and Autonomous Driving, pp. 1519–1535. Springer London, 2012.

[33] U.S. National Highway Traffic Safety Administration, "Preliminary statement of policy concerning automated vehicles," 2013.

[34] SAE International, "J3016 :taxonomy and definitions for terms related to on-road motor vehicle automated driving systems," 2014.

[35] International Organization of Motor Vehicle Manufacturers (OICA), "Automated driving," 2015.

[36] K. I. Naka and W. A. H. Rushton, "S-potentials from luminosity units in the retina of fish (Cyprinidae)," *The Journal of Physiology*, vol. 185, no. 3, pp. 536–555, 1966.

[37] D. A. Baylor and M. G. F. Fuortes, "Electrical responses of single cones in the retina of the turtle," *The Journal of Physiology*, vol. 207, pp. 77–92, 1970.

[38] R. M. Boynton and D. N. Whitten, "Visual Adaptation in Monkey Cones: Recordings of Late Receptor Potentials," *Science*, vol. 170, pp. 1423–1426, 1970.

[39] F. S. Werblin, "Adaptation in a vertebrate retina: intracellular recording in Necturus," *Journal of Neurophysiology*, vol. 34, pp. 228–241, 1971.

[40] J. Kleinschmidt and J. E. Dowling, "Intraceltular Recordings from Gecko Photoreceptors during Light and Dark Adaptation," *The Journal of General Physiology*, vol. 66, pp. 617–648, 1975.

[41] D. C. Hood and P. A. Hock, "Light adaptation of the receptors: Increment threshold functions for the frog's rods and cones," *Vision Research*, vol. 15, pp. 545–553, 1975.

[42] S. B. Laughlin and R. C. Hardie, "Common strategies for light adaptation in the peripheral visual systems of fly and dragonfly," *Journal of Comparative Physiology A*, vol. 128, no. 4, pp. 319–340, 1978.

[43] W. S. Geisler, "Effects of bleaching and backgrounds on the flash response of the cone system," *Journal of Physiology*, vol. 50, pp. 413–434, 1981.

[44] L. Michaelis and M. L. Menten, "The Kinetics of Invertase Action (Die Kinetik der Invertinwirkung)," 1913.

[45] G. Svaetichin, "The cone action potential," *Acta Physiologica*, vol. 29, pp. 565–600, 1953.

[46] T. Oikawa, T. Ogawa, and K. Motokawa, "Origin of so-called cone action potential," *Journal of Neurophysiology*, vol. 22, pp. 102–111, 1959.

[47] R. A. Normann and I. Perlman, "The Effects of Background Illumination on the Photoresponses of Red and Green Cones," *The Journal of Physiology*, vol. 286, pp. 491–507, 1979.

[48] T. Matić and S. B. Laughlin, "Changes in the intensity-response function of an insect's photoreceptors due to light adaptation," *Journal of comparative physiology*, vol. 145, no. 2, pp. 169–177, 1981.

[49] Y. Gu, J. Oberwinkler, M. Postma, and R. C. Hardie, "Mechanisms of light adaptation in Drosophila photoreceptors," *Current biology*, vol. 15, no. 13, pp. 1228–34, 2005.

[50] J. M. Valeton, "PHOTORECEPTOR LIGHT ADAPTATION MODELS: AN EVALUATION," *Vision Research*, vol. 23, no. 12, pp. 1549–1554, 1983.

[51] R. M. Steinman, Z. Pizlo, and F. J. Pizlo, "Phi is not beta, and why wertheimer's discovery launched the gestalt revolution," *Vision Research*, vol. 40, no. 17, pp. 2257 – 2264, 2000.

[52] P. Cavanagh and G. Mather, "Motion: The long and short of it," *Spatial vision*, vol. 4, no. 2, pp. 103–129, 1989.

[53] C. Chubb and G. Sperling, "Drift-balanced random stimuli: a general basis for studying non-fourier motion perception," *J. Opt. Soc. Am. A*, vol. 5, no. 11, pp. 1986–2007, 1988.

[54] S. Nishida, T. Ledgeway, and M. Edwards, "Dual multiple-scale processing for motion in the human visual system," *Vision Research*, vol. 37, no. 19, pp. 2685 – 2698, 1997.

[55] S. Ullman, "Analysis of visual motion by biological and computer systems," *Computer*, vol. 14, no. 8, pp. 57–69, 1981.

[56] F. Raudies, "Optic flow." `http://www.scholarpedia.org/article/Optic_flow`.

[57] H. Wallach, "Über visuell wahrgenommene bewegungsrichtung," *Psychologische Forschung*, vol. 20, no. 1, pp. 325–380, 1935.

[58] G. A. Horridge, "The compound eye of insects," *Scientific American*, 1977.

[59] N. Franceschini, *Photoreceptor Optics*, ch. Sampling of the Visual Environment by the Compound Eye of the Fly: Fundamentals and Applications, pp. 98–125. Springer Berlin Heidelberg, 1975.

[60] N. Franceschini, "Small brains, smart machines: From fly vision to robot vision and back again," *Proceedings of the IEEE*, vol. 102, no. 5, pp. 751–781, 2014.

[61] M. F. Land, "Visual acuity in insects.," *Annual review of entomology*, vol. 42, no. 46, pp. 147–177, 1997.

[62] N. Franceschini, A. Riehle, and A. Nestour, *Facets of vision*, ch. Directionally selective motion detection by insect neurons, pp. 360–390. 1989.

[63] E. Warrant, "Vision in the dimmest habitats on earth," *Journal of Comparative Physiology A*, vol. 190, no. 10, pp. 765–789, 2004.

[64] M. F. Land, *Handbook of Sensory Physiology, Vol. VII/6B*, ch. Optics and Vision in Invertebrates, pp. 472–592. Springer, 1981.

[65] D. G. Stavenga, *Invertebrate Vision*, ch. Invertebrate Photoreceptor Optics, pp. 1–42. Cambridge University Press, 2006.

[66] K. Kirschfeld, *Neural Principles in Vision*, ch. The Resolution of Lens and Compound Eyes, pp. 354–370. Springer Berlin Heidelberg, 1976.

[67] J. K. Douglass and N. J. Strausfeld, "Visual motion-detection circuits in flies: Parallel direction- and non-direction-sensitive pathways between the medulla and lobula plate," *The Journal of Neuroscience*, vol. 16, no. 15, pp. 4551–4562, 1996.

[68] H. G. Krapp and R. Hengstenberg, "Estimation of self-motion by optic flow processing in single visual interneurons," *Nature*, vol. 384, no. 6608, pp. 463–466, 1996.

[69] H. G. Krapp, B. Hengstenberg, and R. Hengstenberg, "Dendritic structure and receptive-field organization of optic flow processing interneurons in the fly," *Journal of Neurophysiology*, vol. 79, no. 4, pp. 1902–1917, 1998.

[70] K. Hausen, "Motion sensitive interneurons in the optomotor system of the fly," *Biological Cybernetics*, vol. 45, no. 2, pp. 143–156, 1982.

[71] R. Hengstenberg, K. Hausen, and B. Hengstenberg, "The number and structure of giant vertical cells (vs) in the lobula plate of the blowfly calliphora erythrocephala," *Journal of Comparative Physiology A*, vol. 149, pp. 163–177, 1982.

[72] M. R. Ibbotson, "Wide-field motion-sensitive neurons tuned to horizontal movement in the honeybee, apis mellifera," *Journal of Comparative Physiology A*, vol. 168, no. 1, pp. 91–102, 1991.

[73] M. Ibbotson, "Evidence for velocity–tuned motion-sensitive descending neurons in the honeybee," *Proceedings of the Royal Society of London B: Biological Sciences*, vol. 268, no. 1482, pp. 2195–2201, 2001.

[74] M. Weckstrom, "Function of compound eye." `http://www.scholarpedia.org/article/Function_of_compound_eye`.

[75] B. Hassenstein and W. Reichardt, "Systemtheoretische analyse der zeit-, reihenfolgen- und vorzeichenauswertung bei der bewegungsperzeption des rüsselkäfers chlorophanus," *Zeitschrift für Naturforschung B*, vol. 11, no. 9, pp. 513–524, 1956.

[76] W. Reichardt, *Sensory Communication*, ch. Autocorrelation, a principle for the evaluation of sensory information by the central nervous system, pp. 303–317. MIT Press, 1961.

[77] M. Egelhaaf and W. Reichardt, "Dynamic response properties of movement detectors: Theoretical analysis and electrophysiological investigation in the visual system of the fly," *Biological Cybernetics*, vol. 56, no. 2, pp. 69–87, 1987.

[78] A. Borst and M. Egelhaaf, "Principles of visual motion detection," *Trends in Neurosciences*, vol. 12, no. 8, pp. 297 – 306, 1989.

[79] A. Borst, "Correlation versus gradient type motion detectors: the pros and cons," *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, vol. 362, no. 1479, pp. 369–374, 2007.

[80] J. P. H. van Santen and G. Sperling, "Elaborated reichardt detectors," *J. Opt. Soc. Am. A*, vol. 2, no. 2, pp. 300–321, 1985.

[81] A. Riehle and N. Franceschini, "Motion detection in flies: Parametric control over on-off pathways," *Experimental Brain Research*, vol. 54, no. 2, pp. 390–394, 1984.

[82] D. F. Reiff, J. Plett, M. Mank, O. Griesbeck, and A. Borst, "Visualizing retinotopic half-wave rectified input to the motion detection circuitry of Drosophila," *Nat Neurosci*, vol. 13, no. 8, pp. 973–978, 2010.

[83] J. Limb and J. Murphy, "Estimating the velocity of moving images in television signals," *Computer Graphics and Image Processing*, vol. 4, no. 4, pp. 311 – 327, 1975.

[84] C. L. Fennema and W. B. Thompson, "Velocity determination in scenes containing several moving objects," *Computer Graphics and Image Processing*, vol. 9, no. 4, pp. 301 – 315, 1979.

[85] E. C. Hildreth and C. Koch, "The analysis of visual motion: From computational theory to neuronal mechanisms," *Annual Review of Neuroscience*, vol. 10, no. 1, pp. 477–533, 1987.

[86] M. V. Srinivasan, "Generalized gradient schemes for the measurement of two-dimensional image motion," *Biological Cybernetics*, vol. 63, no. 6, pp. 421–431, 1990.

[87] M. Potters and W. Bialek, "Statistical mechanics and visual signal processing," *Journal de Physique I*, vol. 4, no. 11, pp. 1755–1775, 1994.

[88] J. Haag, W. Denk, and A. Borst, "Fly motion vision is based on reichardt detectors regardless of the signal-to-noise ratio," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 46, pp. 16333–16338, 2004.

[89] J. J. Gibson, "The perception of the visual world," *The American Journal of Psychology*, vol. 64, no. 3, pp. 440–444, 1951.

[90] v. H. Helmholtz, "Treatise on physiological optics," *Science*, vol. 61, no. 1574, pp. 235–236, 1925.

[91] M. V. Srinivasan, M. Lehrer, S. W. Zhang, and G. A. Horridge, "How honeybees measure their distance from objects of unknown size," *Journal of Comparative Physiology A*, vol. 165, no. 5, pp. 605–613, 1989.

[92] M. Srinivasan, S. Zhang, M. Lehrer, and T. Collett, "Honeybee navigation en route to the goal: visual flight control and odometry," *Journal of Experimental Biology*, vol. 199, no. 1, pp. 237–244, 1996.

[93] B. Ronacher and R. Wehner, "Desert ants cataglyphis fortis use self-induced optic flow to measure distances travelled," *Journal of Comparative Physiology A*, vol. 177, no. 1, pp. 21–27, 1995.

[94] W. H. Warren, B. A. Kay, W. D. Zosh, A. P. Duchon, and S. Sahuc, "Optic flow is used to control human walking," *Nat Neurosci*, vol. 4, no. 2, pp. 213–216, 2001.

[95] D. A. Gordon, "Static and dynamic visual fields in human space perception," *J. Opt. Soc. Am.*, vol. 55, no. 10, pp. 1296–1303, 1965.

[96] J. J. Koenderink and A. J. van Doorn, "Local structure of movement parallax of the plane," *J. Opt. Soc. Am.*, vol. 66, no. 7, pp. 717–723, 1976.

[97] H. C. Longuet-Higgins and K. Prazdny, "The interpretation of a moving retinal image," *Proceedings of the Royal Society of London B: Biological Sciences*, vol. 208, no. 1173, pp. 385–397, 1980.

[98] B. D. Lucas and T. Kanade, "An Iterative Image Registration Technique with an Application to Stereo Vision," in *Proceedings of Imaging Understanding Workshop*, vol. 130, pp. 121–130, 1981.

[99] B. K. P. Horn and B. G. Schunck, "Determining optical flow," *Artificial Intelligence*, vol. 17, pp. 185–203, 1981.

[100] J. J. Koenderink and A. J. V. Doorn, "Facts on optic flow," *Biological Cybernetics*, vol. 56, no. 4, pp. 247–254, 1987.

[101] R. Möckel, *Bio-Inspired Optical Flow Vision Sensors for Visual Guidance of Autonomous Robots*. PhD thesis, ETH Zurich, 2012.

[102] H. Zimmer, A. Bruhn, J. Weickert, L. Valgaerts, A. Salgado, B. Rosenhahn, and H.-P. Seidel, *Energy Minimization Methods in Computer Vision and Pattern Recognition: 7th International Conference, EMMCVPR 2009, Bonn, Germany, August 24-27, 2009. Proceedings*, ch. Complementary Optic Flow, pp. 207–220. Springer Berlin Heidelberg, 2009.

[103] A. Bab-Hadiashar and D. Suter, "Robust optic flow computation," *International Journal of Computer Vision*, vol. 29, no. 1, pp. 59–77, 1998.

[104] C. Mota, I. Stuke, and E. Barth, "Analytical solutions for multiple motions," in *Proceedings of IEEE International Conference on Image Processing*, 2001.

[105] J.-Y. Bouguet, "Pyramidal implementation of the lucas kanade feature tracker," *Intel Corporation, Microprocessor Research Labs*, 2000.

[106] A. Kumar, A. R. Tannenbaum, and G. J. Balas, "Optical flow: a curve evolution approach," *IEEE Transactions on Image Processing*, vol. 5, no. 4, pp. 598–610, 1996.

[107] A. Bruhn, J. Weickert, T. Kohlberger, and C. Schnörr, "A multigrid platform for real-time motion computation with discontinuity-preserving variational methods," *International Journal of Computer Vision*, vol. 70, no. 3, pp. 257–277.

[108] A. Bruhn, J. Weickert, and C. Schnörr, "Lucas/kanade meets horn/schunck: Combining local and global optic flow methods," *International Journal of Computer Vision*, vol. 61, no. 3, pp. 211–231, 2005.

[109] S. S. Beauchemin and J. L. Barron, "The computation of optical flow," *ACM Comput. Surv.*, vol. 27, no. 3, pp. 433–466, 1995.

[110] L. Chen, H. Yang, T. Takaki, and I. Ishii, "Real-time optical flow estimation using multiple frame-straddling intervals," *Journal of Robotics and Mechatronics*, vol. 24, no. 4, pp. 686–698, 2012.

[111] M. V. Srinivasan, "An image-interpolation technique for the computation of optic flow and egomotion," *Biological Cybernetics*, vol. 71, no. 5, pp. 401–415, 1994.

[112] E. H. Adelson and J. R. Bergen, "Spatiotemporal energy models for the perception of motion," *J. Opt. Soc. Am. A*, vol. 2, no. 2, pp. 284–299, 1985.

[113] D. J. Fleet and A. D. Jepson, "Computation of component image velocity from local phase information," *International Journal of Computer Vision*, vol. 5, no. 1, pp. 77–104, 1990.

[114] T. Delbruck, "Silicon retina with correlation-based, velocity-tuned pixels," *IEEE Transactions on Neural Networks*, vol. 4, no. 3, pp. 529–541, 1993.

[115] M. V. Srinivasan, M. Poteser, and K. Kral, "Motion detection in insect orientation and navigation," *Vision Research*, vol. 39, no. 16, pp. 2749 – 2766, 1999.

[116] H. Bulthoff, J. Little, and T. Poggio, "A parallel algorithm for real-time computation of optical flow," *Nature*, vol. 337, no. 6207, pp. 549–553, 1989.

[117] C. Blanes, *Appareil Visuel Elémentaire pour la Navigation à vue d'un Robot Mobile Autonome*. PhD thesis, Univ. Aix-Marseille II, Marseille, France, 1986.

[118] J.-M. Pichon, C. Blanes, and N. Franceschini, "Visual guidance of a mobile robot equipped with a network of self-motion sensors," *Proc. SPIE*, vol. 1195, pp. 44–55, 1989.

[119] N. Franceschini, J. M. Pichon, and C. Blanes, "From insect vision to robot vision," *Phil. Trans. Royal Soc. London B*, vol. 337, pp. 283–294, 1992.

[120] R. G. Benson and T. Delbrück, "Direction selective silicon retina that uses null inhibition," in *Advances in Neural Information Processing Systems 4* (J. E. Moody, S. J. Hanson, and R. P. Lippmann, eds.), pp. 756–763, Morgan-Kaufmann, 1992.

[121] R. Moeckel and S. C. Liu, "Motion detection circuits for a time-to-travel algorithm," in *2007 IEEE International Symposium on Circuits and Systems*, pp. 3079–3082, 2007.

[122] F. Ruffier, S. Viollet, S. Amic, and N. Franceschini, "Bio-inspired optical flow circuits for the visual guidance of micro air vehicles," in *Proc. IEEE Int. Symp. Circuits Syst.*, 2003.

[123] J. S. Albus and T. H. Hong, "Motion, depth, and image flow," in *Robotics and Automation, 1990. Proceedings., 1990 IEEE International Conference on*, vol. 2, pp. 1161–1170, 1990.

[124] H. Bruggeman, W. Zosh, and W. Warren, "Optic flow drives human visuo-locomotor adaptation," *Current Biology*, vol. 17, no. 23, pp. 2035 – 2040, 2007.

[125] D. R. W. Wylie, R. G. Glover, and J. D. Aitchison, "Optic flow input to the hippocampal formation from the accessory optic system," *The Journal of Neuroscience*, vol. 19, no. 13, pp. 5514–5527, 1999.

[126] E. Baird, M. V. Srinivasan, S. Zhang, R. Lamont, and A. Cowling, *From Animals to Animats*, ch. Visual Control of Flight Speed and Height in the Honeybee, pp. 40–51. Springer Berlin Heidelberg, 2006.

[127] B. Hassenstein, "Ommatidienraster und afferente bewegungsintegration," *Zeitschrift für vergleichende Physiologie*, vol. 33, no. 4, pp. 301–326, 1951.

[128] R. Hengstenberg, "Mechanosensory control of compensatory head roll during flight in the blowfly-calliphora erythrocephala meig.," *Journal of Comparative Physiology A*, vol. 163, no. 2, pp. 151–165, 1988.

[129] R. Hengstenberg, "Pitch control of head posture in walking and flying drosophila," in *Proceeding of Invertebrate Vision Conference*, 2001.

[130] J. Hateren and C. Schilstra, "Blowfly flight and optic flow. ii. head movements during flight," *Journal of Experimental Biology*, vol. 202, no. 11, pp. 1491–1500, 1999.

[131] J. S. Kennedy, "The visual responses of flying mosquitoes.," *Proceedings of the Zoological Society of London*, vol. A109, no. 4, pp. 221–242, 1940.

[132] J. S. Kennedy, "The migration of the desert locust (schistocerca gregaria forsk.). i. the behaviour of swarms. ii. a theory of long-range migrations," *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, vol. 235, no. 625, pp. 163–290, 1951.

[133] S. N. Fry, N. Rohrseitz, A. D. Straw, and M. H. Dickinson, "Visual control of flight speed in drosophila melanogaster," *Journal of Experimental Biology*, vol. 212, no. 8, pp. 1120–1130, 2009.

[134] E. Baird, T. Kornfeldt, and M. Dacke, "Minimum viewing angle for visually guided ground speed control in bumblebees," *Journal of Experimental Biology*, vol. 213, no. 10, pp. 1625–1632, 2010.

[135] J. R. Serres, G. P. Masson, F. Ruffier, and N. Franceschini, "A bee in the corridor: centering and wall-following," *Naturwissenschaften*, vol. 95, no. 12, pp. 1181–1187, 2008.

[136] G. Portelli, F. Ruffier, F. L. Roubieu, and N. Franceschini, "Honeybees' speed depends on dorsal as well as lateral, ventral and frontal optic flows," *PLoS ONE*, vol. 6, no. 5, pp. 1–10, 2011.

[137] P. Bhagavatula, C. Claudianos, M. Ibbotson, and M. Srinivasan, "Optic flow cues guide flight in birds," *Current Biology*, vol. 21, no. 21, pp. 1794 – 1799, 2011.

[138] W. K. Page and C. J. Duffy, "Cortical neuronal responses to optic flow are shaped by visual strategies for steering," *Cerebral Cortex*, vol. 18, no. 4, pp. 727–739, 2008.

[139] H. Bruggeman and W. H. Warren, "The Direction of Walking - but Not Throwing or Kicking - Is Adapted by Optic Flow," *Psychological science*, vol. 21, no. 7, pp. 1006–1013, 2010.

[140] G. K. Kountouriotis, C. D. Mole, N. Merat, and R. M. Wilkie, "The need for speed: global optic flow speed influences steering," *Royal Society Open Science*, vol. 3, no. 5, 2016.

[141] B. J. Correia Grácio, J. E. Bos, M. M. Paassen, and M. Mulder, "Perceptual scaling of visual and inertial cues," *Experimental Brain Research*, vol. 232, no. 2, pp. 637–646, 2013.

[142] C. Authié, *Controle visuel du déplacement en trajectoire courbe : approche sensorimotrice du role structurant du flux optique*. PhD thesis, Aix-Marseille University, 2011.

[143] W. H. Warren and B. R. Fajen, *Optic Flow and Beyond*, ch. From Optic Flow to Laws of Control, pp. 307–337. Norwell, MA, USA: Kluwer Academic Publishers, 2004.

[144] K. V. Frisch, *The Dance Language and Orientation of Bees*. Harvard University Press, 1967.

[145] H. E. Esch and J. E. Burns, "Honeybees use optic flow to measure the distance of a food source," *Naturwissenschaften*, vol. 82, no. 1, pp. 38–40, 1995.

[146] M. Srinivasan, S. Zhang, and N. Bidwell, "Visually mediated odometry in honeybees," *Journal of Experimental Biology*, vol. 200, no. 19, pp. 2513–2522, 1997.

[147] M. A. Eckles, D. W. Roubik, and J. C. Nieh, "A stingless bee can use visual odometry to estimate both height and distance," *Journal of Experimental Biology*, vol. 215, no. 18, pp. 3155–3160, 2012.

[148] J. Tautz, S. Zhang, J. Spaethe, A. Brockmann, A. Si, and M. Srinivasan, "Honeybee odometry: Performance in varying natural terrain," *PLoS Biol*, vol. 2, 07 2004.

[149] M. Dacke and M. V. Srinivasan, "Two odometers in honeybees?," *Journal of Experimental Biology*, vol. 211, no. 20, pp. 3281–3286, 2008.

[150] B. Ronacher, K. Gallizzi, S. Wohlgemuth, and R. Wehner, "Lateral optic flow does not influence distance estimation in the desert ant cataglyphis fortis," *Journal of Experimental Biology*, vol. 203, no. 7, pp. 1113–1121, 2000.

[151] H. Wolf, "Odometry and insect navigation," *Journal of Experimental Biology*, vol. 214, no. 10, pp. 1629–1641, 2011.

[152] A. Barron and M. V. Srinivasan, "Visual regulation of ground speed and headwind compensation in freely flying honey bees (apis mellifera l.)," *Journal of Experimental Biology*, vol. 209, no. 5, pp. 978–984, 2006.

[153] G. Portelli, F. Ruffier, and N. Franceschini, "Honeybees change their height to restore their optic flow," *Journal of Comparative Physiology A*, vol. 196, no. 4, pp. 307–313, 2010.

[154] V. M. Srinivasan, W. S. Zhang, S. J. Chahl, E. Barth, and S. Venkatesh, "How honeybees make grazing landings on flat surfaces," *Biological Cybernetics*, vol. 83, no. 3, pp. 171–183, 2000.

[155] F. Expert, *Flying robot inspired by insects: From optic flow sensing to visually guided strategies to control a Micro Aerial Vehicle*. PhD thesis, Aix-Marseille University, 2013.

[156] A. Spivak, A. Belenky, A. Fish, and O. Yadid-Pecht, "Wide-Dynamic-Range CMOS Image Sensors - Comparative Performance Analysis," *IEEE Transactions on Electron Devices*, vol. 56, no. 11, pp. 2446–2461, 2009.

[157] S. Kavadias, B. Dierickx, D. Scheffer, A. Alaerts, D. Uwaerts, and J. Bogaerts, "A logarithmic response cmos image sensor with on-chip calibration," *IEEE Journal of Solid-State Circuits*, vol. 35, no. 8, pp. 1146–1152, 2000.

[158] E. Labonne, G. Sicard, M. Renaudin, and P. D. Berger, "A 100db dynamic range cmos image sensor with global shutter," in *2006 13th IEEE International Conference on Electronics, Circuits and Systems*, pp. 1133–1136, 2006.

[159] J. Fox, C. E.and Hynecek and D. R. Dykaar, "Wide-dynamic-range pixel with combined linear and logarithmic response and increased signal swing," *Proc. SPIE*, 2006.

[160] G. Storm, R. Henderson, J. E. D. Hurwitz, D. Renshaw, K. Findlater, and M. Purcell, "Extended dynamic range from a combined linear-logarithmic cmos image sensor," *IEEE Journal of Solid-State Circuits*, vol. 41, no. 9, pp. 2095–2106, 2006.

[161] S. Decker, D. McGrath, K. Brehmer, and C. G. Sodini, "A 256x256 cmos imaging array with wide dynamic range pixels and column-parallel digital output," *IEEE Journal of Solid-State Circuits*, vol. 33, no. 12, pp. 2081–2091, 1998.

[162] D. Hertel, A. Betts, R. Hicks, and M. ten Brinke, "An adaptive multiple-reset cmos wide dynamic range imager for automotive vision applications," in *Intelligent Vehicles Symposium, 2008 IEEE*, pp. 614–619, 2008.

[163] K. P. Frohmader, "A novel mos compatible light intensity-to-frequency converter suited for monolithic integration," *IEEE Journal of Solid-State Circuits*, vol. 17, no. 3, pp. 588–591, 1982.

[164] E. Culurciello, R. Etienne-Cummings, and K. A. Boahen, "A biomorphic digital image sensor," *IEEE Journal of Solid-State Circuits*, vol. 38, no. 2, pp. 281–294, 2003.

[165] A. Kitchen, A. Bermak, and A. Bouzerdoum, "Pwm digital pixel sensor based on asynchronous self-resetting scheme," *IEEE Electron Device Letters*, vol. 25, no. 7, pp. 471–473, 2004.

[166] C. Shoushun and A. Bermak, "Arbitrated time-to-first spike cmos image sensor with on-chip histogram equalization," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 15, no. 3, pp. 346–357, 2007.

[167] M. Mase, S. Kawahito, M. Sasaki, Y. Wakamori, and M. Furuta, "A wide dynamic range cmos image sensor with multiple exposure-time signal outputs and 12-bit column-parallel cyclic a/d converters," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 12, pp. 2787–2795, 2005.

[168] O. Yadid-Pecht and A. Belenky, "In-pixel autoexposure cmos aps," *IEEE Journal of Solid-State Circuits*, vol. 38, no. 8, pp. 1425–1428, 2003.

[169] J. Carneiro, S.-H. Ieng, C. Posch, and R. Benosman, "Event-based 3D reconstruction from neuromorphic retinas," *Neural Networks*, vol. 45, pp. 27–38, 2013.

[170] C. Pacoret and S. Régnier, "A review of haptic optical tweezers for an interactive microworld exploration," *Rev. Sci. Instrum.*, vol. 84, 2013.

[171] C. A. Mead and M. Mahowald, "A silicon model of early visual processing," *Neural Networks*, vol. 1, pp. 91–97, 1988.

[172] T. Delbruck and C. A. Mead, "An electronic photoreceptor sensitive to small changes in intensity," in *Advances in Neural Information Processing Systems*, pp. 720–726, 1989.

[173] F. S. Werblin, "Control of Retinal Sensitivity II: Lateral Interactions at the Outer Plexiform Layer," *The Journal of General Physiology*, vol. 63, no. 1, pp. 62–87, 1974.

[174] K. A. Boahen and A. G. Andreou, "A Contrast Sensitive Silicon Retina with Reciprocal Synapses," in *Advances in neural information processing systems*, pp. 764–772, Morgan Kaufmann, 1991.

[175] T. Delbrück and C. A. Mead, "Analog VLSI Adaptive , Logarithmic , Wide-Dynamic-Range Photoreceptor," in *ISCAS*, pp. 339–342, 1994.

[176] T. Delbrück, *Investigations of visual transduction and motion processing*. PhD thesis, California Inst. Tech., Pasadena, California, 1993.

[177] F. Expert, S. Viollet, and F. Ruffier, "Outdoor field performances of insect-based visual motion sensors," *J. Field Robot.*, vol. 28, no. 4, pp. 529–541, 2011.

[178] G. Sicard, H. Abbas, H. Amhaz, H. Zimouche, R. Rolland, and D. Alleysson, "A cmos hdr imager with an analog local adaptation," in *Int. Image Sensor Workshop (IISW '13)*, pp. 1–4, 2013.

[179] L. Meylan, D. Alleysson, and S. Süsstrunk, "Model of retinal local adaptation for the tone mapping of color filter array images," *Journal of the Optical Society of America. A, Optics, image science, and vision*, vol. 24, no. 9, pp. 2807–2816, 2007.

[180] S.-C. Liu, T. Delbruck, G. Indiveri, A. Whatley, and R. Douglas, *Event-Based Neuromorphic Systems*. Wiley, 2014.

[181] K. A. Zaghloul and K. Boahen, "Optic nerve signals in a neuromorphic chip i: Outer and inner retina models," *IEEE Trans. Bio-Med. Eng*, pp. 657–666, 2004.

[182] P. F. Ruedi, P. Heim, F. Kaess, E. Grenet, F. Heitger, P. Y. Burgi, S. Gyger, and P. Nussbaum, "A 128x128 pixel 120 db dynamic range vision sensor chip for image contrast and orientation extraction," in *Solid-State Circuits Conference, 2003. Digest of Technical Papers. ISSCC. 2003 IEEE International*, pp. 226–490, 2003.

[183] P. Lichtsteiner, C. Posch, and T. Delbruck, "A 128x128 120 db 15 μs latency asynchronous temporal contrast vision sensor," *Solid-State Circuits, IEEE Journal of*, vol. 43, no. 2, pp. 566–576, 2008.

[184] C. Posch, D. Matolin, and R. Wohlgenannt, "A QVGA 143 dB Dynamic Range Frame-Free PWM Image Sensor With Lossless Pixel-Level Video Compression and Time-Domain CDS," *IEEE Journal of Solid-State Circuits*, vol. 46, no. 1, pp. 259–275, 2011.

[185] A. N. Belbachir, M. Litzenberger, S. Schraml, M. Hofstätter, D. Bauer, P. Schön, M. Humenberger, C. Sulzbachner, T. Lunden, and M. Merne, "Care: A dynamic stereo vision sensor system for fall detection," in *2012 IEEE International Symposium on Circuits and Systems*, pp. 731–734, 2012.

[186] C. Posch, "Bio-inspired vision," *Journal of Instrumentation*, vol. 7, no. 01, p. C01054, 2012.

[187] R. Berner, C. Brandli, M. Yang, S. C. Liu, and T. Delbruck, "A 240x180 10mw 12us latency sparse-output vision sensor for mobile applications," in *VLSI Circuits (VLSIC), 2013 Symposium on*, pp. C186–C187, 2013.

[188] C. Brandli, R. Berner, M. Yang, S. C. Liu, and T. Delbruck, "A 240x180 130db 3us latency global shutter spatiotemporal vision sensor," *IEEE Journal of Solid-State Circuits*, vol. 49, no. 10, pp. 2333–2341, 2014.

[189] J. Diaz, E. Ros, R. Agis, and J. L. Bernier, "Superpipelined high-performance optical-flow computation architecture," *Computer Vision and Image Understanding*, vol. 112, no. 3, pp. 262 – 273, 2008.

[190] D. Watman and H. Murayama, "Design of a miniature, multi-directional optical flow sensor for micro aerial vehicles," in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, pp. 2986–2991, 2011.

[191] J. Plett, A. Bahl, M. Buss, K. Kühnlenz, and A. Borst, "Bio-inspired visual ego-rotation sensor for mavs," *Biological Cybernetics*, vol. 106, no. 1, pp. 51–63, 2012.

[192] D. Honegger, L. Meier, P. Tanskanen, and M. Pollefeys, "An open source and open hardware embedded metric optical flow cmos camera for indoor and outdoor applications," in *Robotics and Automation (ICRA), 2013 IEEE International Conference on*, pp. 1736–1741, 2013.

[193] R. Lyon and M. Haeberli, "Designing and testing the optical mouse," *VLSI Design*, vol. 1, pp. 20–29, 1982.

[194] R. Chan, A. Mulla, and K. Stol, "Characterisation of low-cost optical flow sensors," in *Australasian Conference on Robotics and Automation (ACRA)*, 2010.

[195] D. Font, M. Tresanchez, T. Pallejà, M. Teixidó, and J. Palacín, "Characterization of a Low-Cost Optical Flow Sensor When Using an External Laser as a Direct Illumination Source," *Sensors (Basel, Switzerland)*, vol. 11, no. 12, pp. 11856–11870, 2011.

[196] J. D. Jackson, D. W. Callahan, and J. Marstrander, "A rationale for the use of optical mice chips for economic and accurate vehicle tracking," in *Automation Science and Engineering, 2007. CASE 2007. IEEE International Conference on*, pp. 939–944, 2007.

[197] J. Tanner and C. Mead, *VLSI Signal Processing II*, ch. An Integrated Analog Optical Motion Sensor, pp. 59–76. IEEE Press, 1986.

[198] A. J. Moore and C. Koch, "Multiplication-based analog motion detection chip," *Proc. SPIE*, vol. 1473, pp. 66–75, 1991.

[199] A. A. Stocker and R. J. Douglas, "Analog integrated 2-d optical flow sensor with programmable pixels," in *Circuits and Systems, 2004. ISCAS '04. Proceedings of the 2004 International Symposium on*, vol. 3, pp. 9–12, 2004.

[200] A. A. Stocker, "Analog integrated 2-d optical flow sensor," *Analog Integrated Circuits and Signal Processing*, vol. 46, no. 2, pp. 121–138, 2006.

[201] R. R. Harrison and C. Koch, "A robust analog vlsi reichardt motion sensor," *Analog Integrated Circuits and Signal Processing*, vol. 24, no. 3, pp. 213–229, 2000.

[202] C. M. Higgins and S. A. Shams, "A biologically inspired modular vlsi system for visual measurement of self-motion," *IEEE Sensors Journal*, vol. 2, no. 6, pp. 508–528, 2002.

[203] P. Xu, J. S. Humbert, and P. Abshire, "Analog vlsi implementation of wide-field integration methods," *Journal of Intelligent & Robotic Systems*, vol. 64, no. 3, pp. 465–487, 2011.

[204] R. Sarpeshkar, J. Kramer, G. Indiveri, and C. Koch, "Analog vlsi architectures for motion processing: from fundamental limits to system applications," *Proceedings of the IEEE*, vol. 84, no. 7, pp. 969–987, 1996.

[205] A. A. Stocker, *Analog VLSI Circuits for the Perception of Visual Motion*. Wiley, 2006.

[206] J. Kramer, R. Sarpeshkar, and C. Koch, "An analog vlsi velocity sensor," in *Circuits and Systems, 1995. ISCAS '95., 1995 IEEE International Symposium on*, vol. 1, pp. 413–416, 1995.

[207] R. Deutschmann, *Analog VLSI Motion Sensors*. PhD thesis, Technische Universität München (Germany) and California Institute of Technology (USA), 1997.

[208] G. L. Barrows and C. Neely, "Mixed-mode vlsi optic flow sensors for in-flight control of a micro air vehicle," *Proc. SPIE*, vol. 4109, pp. 52–63, 2000.

[209] R. Benosman, C. Clercq, X. Lagorce, S.-H. Ieng, and C. Bartolozzi, "Event-based visual flow," *Neural Networks and Learning Systems, IEEE Transactions on*, vol. 25, no. 2, pp. 407–417, 2014.

[210] M. Giulioni, X. Lagorce, F. Galluppi, and R. B. Benosman, "Event-based computation of motion flow on a neuromorphic analog neural platform," *Frontiers in Neuroscience*, vol. 10, no. 35, 2016.

[211] C. Blanes, *Guidage visuel d'un robot mobile autonome d'inspitation bionique*. PhD thesis, Polytechnical Institute, Grenoble, France, 1991.

[212] M. Pudas, S. Viollet, F. Ruffier, A. Kruusing, S. Amic, S. Leppävuori, and N. Franceschini, "A miniature bio-inspired optic flow sensor based on low temperature co-fired ceramics (ltcc) technology," *Sensors and Actuators A: Physical*, vol. 133, no. 1, pp. 88 – 95, 2007.

[213] F. Aubépart and N. Franceschini, "Bio-inspired optic flow sensors based on fpga: Application to micro-air-vehicles," *Microprocessors and Microsystems*, vol. 31, no. 6, pp. 408 – 419, 2007.

[214] F. Roubieu, F. Expert, G. Sabiron, and F. Ruffier, "A two-directional 1-gram visual motion sensor inspired by the fly's eye," *IEEE Sensor Journal*, vol. 13, pp. 1025–1035, 2013.

[215] F. Expert, F. L. Roubieu, and F. Ruffier, "Interpolation based "time of travel" scheme in a visual motion sensor using a small 2d retina," in *Sensors, 2012 IEEE*, pp. 1–4, 2012.

[216] G. Sabiron, P. Chavent, T. Raharijaona, P. Fabiani, , and F. Ruffier, "Low-speed optic-flow sensor onboard an unmanned helicopter flying outside over fields," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2013.

[217] L. Kerhuel, *Capteurs optiques minimalistes et réflexes oculomoteurs biomimétiques. Applications a la robotique aérienne*. PhD thesis, University of Montpellier, France, 2009.

[218] Y. M. Song, Y. Xie, V. Malyarchuk, J. Xiao, I. Jung, K.-J. Choi, Z. Liu, H. Park, C. Lu, R.-H. Kim, R. Li, K. B. Crozier, Y. Huang, and J. A. Rogers, "Digital cameras with designs inspired by the arthropod eye," *Nature*, vol. 497, no. 7447, pp. 95–99, 2013.

[219] D. Floreano, R. Pericet-Camara, S. Viollet, F. Ruffier, A. Brückner, R. Leitel, W. Buss, M. Menouni, F. Expert, R. Juston, M. K. Dobrzynski, G. L'Eplattenier, F. Recktenwald, H. a. Mallot, and N. Franceschini, "Miniature curved artificial compound eyes," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 110, no. 23, pp. 9267–72, 2013.

[220] A. J. Ijspeert, A. Crespi, D. Ryczko, and J.-M. Cabelguen, "From swimming to walking with a salamander robot driven by a spinal cord model," *Science*, vol. 315, no. 5817, pp. 1416–1420, 2007.

[221] A. D. Marchese, C. D. Onal, and D. Rus, "Autonomous soft robotic fish capable of escape maneuvers using fluidic elastomer actuators," *Soft Robotics*, 2014.

[222] R. J. Wood, "The first takeoff of a biologically inspired at-scale robotic insect," *IEEE Transactions on Robotics*, vol. 24, no. 2, pp. 341–347, 2008.

[223] A. Crespi, K. Karakasiliotis, A. Guignard, and A. J. Ijspeert, "Salamandra robotica ii: An amphibious robot to study salamander-like swimming and walking gaits," *IEEE Transactions on Robotics*, vol. 29, no. 2, pp. 308–320, 2013.

[224] J. K. Hopkins, B. W. Spranklin, and S. K. Gupta, "A survey of snake-inspired robot designs," *Bioinspiration & Biomimetics*, vol. 4, no. 2, p. 021001, 2009.

[225] K. Y. Ma, P. Chirarattananon, S. B. Fuller, and R. J. Wood, "Controlled flight of a biologically inspired, insect-scale robot," *Science*, vol. 340, no. 6132, pp. 603–607, 2013.

[226] A. A. Argyros, D. P. Tsakiris, and C. Groyer, "Biomimetic centering behavior [mobile robots with panoramic sensors]," *IEEE Robotics Automation Magazine*, vol. 11, no. 4, pp. 21–30, 2004.

[227] R. R. Harrison and C. Koch, "A robust analog vlsi motion sensor based on the visual system of the fly," *Auton. Robots*, vol. 7, no. 3, pp. 211–224, 1999.

[228] H. Dahmen, A. Millers, and H. A. Mallot, *Flying Insects and Robots*, ch. Insect-Inspired Odometry by Optic Flow Recorded with Optical Mouse Chips, pp. 115–126. Springer Berlin Heidelberg, 2010.

[229] J. Santos-Victor, G. Sandini, F. Curotto, and S. Garibaldi, "Divergent stereo in autonomous navigation: From bees to robots," *International Journal of Computer Vision*, vol. 14, no. 2, pp. 159–177, 1995.

[230] S. van der Zwaan and J. Santos-Victor, "An insect inspired visual sensor for the autonomous navigation of a mobile robot," in *International Symposium on Intelligent Robotic Systems (SIRS)*, 1999.

[231] J. S. Humbert, J. K. Conroy, C. W. Neely, and G. Barrows, *Flying Insects and Robots*, ch. Wide-Field Integration Methods for Visuomotor Control, pp. 63–71. Springer Berlin Heidelberg, 2010.

[232] J. S. Humbert and A. M. Hyslop, "Bioinspired visuomotor convergence," *IEEE Transactions on Robotics*, vol. 26, no. 1, pp. 121–130, 2010.

[233] A. M. Hyslop and J. S. Humbert, "Autonomous Navigation in Three-Dimensional Urban Environments Using Wide-Field Integration of Optic Flow," *Journal of Guidance Control and Dynamics*, vol. 33, pp. 147–159, 2010.

[234] J. Conroy, G. Gremillion, B. Ranganathan, and J. S. Humbert, "Implementation of wide-field integration of optic flow for autonomous quadrotor navigation," *Autonomous Robots*, vol. 27, no. 3, pp. 189–198, 2009.

[235] T. Netter and N. H. Franceschini, "Neuromorphic optical flow sensing for nap-of-the-earth flight," *Proc. SPIE*, vol. 3838, pp. 208–216, 1999.

[236] T. Netter and N. Francheschini, "A robotic aircraft that follows terrain using a neuromorphic eye," in *Intelligent Robots and Systems, 2002. IEEE/RSJ International Conference on*, vol. 1, pp. 129–134, 2002.

[237] F. Ruffier and N. Franceschini, "Visually guided micro-aerial vehicle: automatic take off, terrain following, landing and wind reaction," in *Robotics and Automation, 2004. Proceedings. ICRA '04. 2004 IEEE International Conference on*, vol. 3, pp. 2339–2346, 2004.

[238] F. Ruffier and N. Franceschini, "Optic flow regulation: the key to aircraft automatic guidance," *Robotics and Autonomous Systems*, vol. 50, no. 4, pp. 177 – 194, 2005.

[239] N. Franceschini, F. Ruffier, and J. Serres, "A bio-inspired flying robot sheds light on insect piloting abilities," *Current Biology*, vol. 17, no. 4, pp. 329 – 335, 2007.

[240] F. Ruffier and N. Franceschini, "Aerial robot piloted in steep relief by optic flow sensors," in *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*, pp. 1266–1273, 2008.

[241] F. L. Roubieu, J. Serres, N. Franceschini, F. Ruffier, and S. Viollet, "A fully-autonomous hovercraft inspired by bees: Wall following and speed control in straight and tapered corridors," in *Robotics and Biomimetics (ROBIO), 2012 IEEE International Conference on*, pp. 1311–1318, 2012.

[242] F. Expert and F. Ruffier, "Flying over uneven moving terrain based on optic-flow cues without any need for reference frames or accelerometers," *Bioinspiration & Biomimetics*, vol. 10, no. 2, p. 026003, 2015.

[243] G. Portelli, J. Serres, F. Ruffier, and N. Franceschini, "Modelling honeybee visual guidance in a 3-d environment," *Journal of Physiology-Paris*, vol. 104, no. 1-2, pp. 27 – 39, 2010.

[244] S. Mota, E. Ros, E. M. Ortigosa, and F. Pelayo, "Bio-inspired motion detection for a blind spot overtaking monitor," *International Journal of Robotics and Automation*, 2004.

[245] R. Laviana, L. Carranza, S. Vargas, G. Linan, and E. Roca, "A bioinspired vision chip architecture for collision detection in automotive applications," *Proc. SPIE*, vol. 5839, pp. 13–24, 2005.

[246] G. Linan-Cembrano, L. Carranza, C. Rind, A. Zarandy, M. Soininen, and A. Rodriguez-Vazquez, "Insect-vision inspired collision warning vision processor for automobiles," *IEEE Circuits and Systems Magazine*, vol. 8, no. 2, pp. 6–24, 2008.

[247] J. Kogler, C. Sulzbachner, and W. Kubinger, *Computer Vision Systems: 7th International Conference on Computer Vision Systems, ICVS 2009 Liège, Belgium, October 13-15, 2009. Proceedings*,

ch. Bio-inspired Stereo Vision System with Silicon Retina Imagers, pp. 174–183. Springer Berlin Heidelberg, 2009.

[248] T. Michalke, R. Kastner, J. Fritsch, and C. Goerick, "Towards a proactive biologically-inspired advanced driver assistance system," in *IEEE Intelligent Vehicles Symposium (IV)*, 2009.

[249] T. Michalke, J. Fritsch, and C. Goerick, "A biologically-inspired vision architecture for resource-constrained intelligent vehicles," *Computer Vision and Image Understanding*, vol. 114, no. 5, pp. 548 – 563, 2010.

[250] A. Morand, *Commande assistée au conducteur basée sur la conduite en formation de type "banc de poissons"*. PhD thesis, University of Bordeaux, 2014.

[251] I. E. Paromtchik and C. Laugier, "Autonomous parallel parking of a nonholonomic vehicle," in *Proc. IEEE Intell. Veh. Symp.*, 1996.

[252] S. Hiramatsu, "Rearview camera based parking assist system with voice guidance." SAE Paper No: 2002-01-0759, 2002.

[253] P. Degerman, J. Pohl, and M. Sethson, "Hough transform for parking space estimation using long range ultrasonic sensors." SAE Paper No: 2006-01-0810, 2006.

[254] W.-J. Park, B.-S. Kim, D.-E. Seo, D.-S. Kim, and K.-H. Lee, "Parking space detection using ultrasonic sensor in parking assistance system," in *Proc. IEEE Intell. Veh. Symp.*, 2008.

[255] T.-h. Hsu, J.-f. Liu, P.-n. Yu, W.-s. Lee, and H. J-s., "Development of an automatic parking system for vehicle," in *Proc. IEEE Veh. Power Prop. Conf.*, 2008.

[256] K. Fintzel and R. Bendahan, "3D parking assistant system," in *Proc. IEEE Intell. Veh. Symp.*, 2004.

[257] J. K. Suhr, H. G. Jung, K. Bae, and J. Kim, "Automatic free parking space detection by using motion stereo-based 3D reconstruction," *Machine Vision and Applications*, vol. 21, pp. 163–176, 2010.

[258] S.-e. Shih, W.-h. Tsai, and S. Member, "A Convenient Vision-Based System for Automatic Detection of Parking Spaces in Indoor Parking Lots Using Wide-Angle Cameras," *IEEE Trans. Intell. Transp. Syst.*, vol. 63, pp. 2521–2532, 2014.

[259] P. Jeevan, F. Harchut, B. Mueller-Bessler, and B. Huhnke, "Realizing Autonomous Valet Parking with automotive grade sensors," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2010.

[260] S. Gorner and H. Rohling, "Parking lot detection with 24 ghz radar sensor," in *3rd Int. Workshop Intell. Transp.*, 2006.

[261] S. Mafrica, S. Godiot, M. Menouni, M. Boyron, F. Expert, R. Juston, N. Marchand, F. Ruffier, and S. Viollet, "A bio-inspired analog silicon retina with michaelis-menten auto-adaptive pixels sensitive to small and large changes in light," *Opt. Express*, vol. 23, no. 5, pp. 5614–5635, 2015.

[262] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, pp. 381–395, 1981.

[263] R. Vidal, "Subspace clustering," *IEEE Signal Processing Magazine*, vol. 28, pp. 52–68, 2011.

[264] A. De Luca, G. Oriolo, and C. Samson, *Feedback control of a nonholonomic car-like robot*. 1998.

[265] L. Liu, J. Wiinschmann, N. P. Aryan, A. Zohny, M. Fischer, S. Kibbel, and A. Rothermel, "An ambient light adaptive subretinal stimulator," in *ESSCIRC '09*, 2009.

[266] D. Drazen, P. Lichtsteiner, P. Häfliger, T. Delbrück, and A. Jensen, "Toward real-time particle tracking using an event-based dynamic vision sensor," *Experiments in Fluids*, vol. 51, no. 5, pp. 1465–1469, 2011.

[267] S. Laughlin and M. Weckstrom, "Fast and slow photoreceptors - a comparative study of the functional diversity of coding and conductances in the Diptera," *Journal of Comparative Physiology A*, vol. 172, no. 5, pp. 593–609, 1993.

[268] K. A. Zaghloul and K. Boahen, "A silicon retina that reproduces signals in the optic nerve," vol. 3, pp. 257–267, 2006.

[269] K. Shimonomura, S. Kameda, A. Iwata, and T. Yagi, "Wide-dynamic-range APS-based silicon retina with brightness constancy," *IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council*, vol. 22, no. 9, pp. 1482–93, 2011.

[270] E. Reinhard and K. Devlin, "Dynamic range reduction inspired by photoreceptor physiology," *IEEE transactions on visualization and computer graphics*, vol. 11, no. 1, pp. 13–24, 2005.

[271] N.-s. Vu, A. Caplier, B. P. St, and H. France, "ILLUMINATION-ROBUST FACE RECOGNITION USING RETINA MODELING," pp. 3289–3292, 2009.

[272] S. Ferradans, M. Bertalmio, E. Provenzi, and V. Caselles, "An analysis of visual adaptation and contrast perception for tone mapping," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 33, no. 10, pp. 2002–2012, 2011.

[273] E. Zrenner, "Will Retinal Implants Restore Vision?," *Nature*, vol. 295, pp. 1022–1025, 2002.

[274] K. Stingl, K. U. Bartz-Schmidt, D. Besch, A. Braun, A. Bruckmann, F. Gekeler, U. Greppmaier, S. Hipp, G. Hörtdörfer, C. Kernstock, H. Koitschev, A. Kusnyerik, H. Sachs, A. Schatz, K. T. Stingl, T. Peters, B. Wilhelm, and E. Zrenner, "Artificial vision with wirelessly powered subretinal electronic implant alpha-ims," *Proc. R. Soc. London, Ser. B*, vol. 280, 2013.

[275] S. Viollet, S. Godiot, R. Leitel, W. Buss, P. Breugnon, M. Menouni, R. Juston, F. Expert, F. Colonnier, G. L'Eplattenier, A. Brückner, F. Kraze, H. Mallot, N. Franceschini, R. Pericet-Camara, F. Ruffier,

and D. Floreano, “Hardware Architecture and Cutting-Edge Assembly Process of a Tiny Curved Compound Eye,” *Sensors*, vol. 110, pp. 21702–21721, 2014.

[276] S.-C. Liu, Kramer, Indiveri, T. Delbruck, and R. Douglas, *Analog VLSI: Circuits and Principles*. MIT Press, 2002.

[277] P. Venier and X. Arreguit, “Réseau de cellules photosensibles et capteur d’images comportant un tel réseau,” 1997.

[278] S. Panzieri, F. Pascucci, and G. Ulivi, “An outdoor navigation system using GPS and inertial platform,” *IEEE/ASME Trans. Mechatronics*, vol. 7, no. 2, pp. 134–142, 2002.

[279] A. Georgiev and P. Allen, “Localisation methods for a mobile robot in urban environments,” *IEEE Trans. Robot. Autom.*, vol. 20, no. 5, pp. 851–864, 2004.

[280] Kistler, “Correvit s-350: 2-axis optical sensors.” `https://www.kistler.com/it/en/applications/automotive-research-test/vehicle-dynamics-durability/dynamics-testing/products/correvit_s_350_2_axis_optical_sensors_cs350_a_`.

[281] G. Tagne, R. Talj, and A. Charara, “Design of passivity-based controllers for lateral dynamics of intelligent vehicles,” in *2015 IEEE Intelligent Vehicles Symposium (IV)*, pp. 1044–1049, 2015.

[282] D. Sun, S. Roth, and M. Black, “Secrets of optical flow estimation and their principles,” in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pp. 2432–2439, 2010.

[283] H. Chao, Y. Gu, and M. Napolitano, “A survey of optical flow techniques for robotics navigation applications,” *Journal of Intelligent & Robotic Systems*, vol. 73, no. 1, pp. 361–372, 2013.

[284] A. Alaeddini and K. A. Morgansen, “Bioinspired navigation for a nonholonomic mobile robot,” *Journal of Aerospace Information Systems*, 2015.

[285] L. R. Garcia Carrillo, I. Fantoni, E. Rondón, and A. Dzul, “Three-dimensional Position and Velocity Regulation of a Quad-Rotorcraft Using Optical Flow,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 51, no. 1, pp. 358–371, 2015.

[286] J. Campbell, R. Sukthankar, and I. Nourbakhsh, “Techniques for evaluating optical flow for visual odometry in extreme terrain,” in *Proc. IEEE Int. Conf. Intell. Robot. Syst.*, pp. 3704–3711, 2004.

[287] P. Corke, D. Strelow, and S. Singh, “Omnidirectional visual odometry for a planetary rover,” in *Intelligent Robots and Systems, 2004. (IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on*, vol. 4, pp. 4007–4012, 2004.

[288] D. Scaramuzza, F. Fraundorfer, and R. Siegwart, “Real-time monocular visual odometry for on-road vehicles with 1-point RANSAC,” in *Proc. IEEE Int. Conf. Robot. Autom.*, pp. 4293–4299, 2009.

[289] F. Fraundorfer and D. Scaramuzza, "Visual odometry: Part ii: Matching, robustness, optimization, and applications," *Robotics Automation Magazine, IEEE*, vol. 19, no. 2, pp. 78–90, 2012.

[290] A. Briod, J.-C. Zufferey, and D. Floreano, "A method for ego-motion estimation in micro-hovering platforms flying in very cluttered environments," *Autonomous Robots*, pp. 1–15, 2015.

[291] C. B. Williams, "Insect migration," *Annual Review of Entomology*, vol. 2, no. 1, pp. 163–180, 1957.

[292] R. B. Srygley and E. G. Oliveira, *Insect movement mechanisms and consequences. Proceedings of the Royal Entomological Society's 20th Symposium*, ch. Orientation mechanisms and migration strategies within the flight boundary layer, pp. 183–206. CABI Publishing, CAB International, 2001.

[293] E. Baird, E. Kreiss, W. Wcislo, E. Warrant, and M. Dacke, "Nocturnal insects use optic flow for flight control," *Biology Letters*, vol. 7, no. 4, pp. 499–501, 2011.

[294] C. Posch, T. Serrano-Gotarredona, B. Linares-Barranco, and T. Delbruck, "Retinomorphic event-based vision sensors: Bioinspired cameras with spiking output," *Proceedings of the IEEE*, vol. 102, no. 10, pp. 1470–1484, 2014.

[295] C. Posch, R. Benosman, and R. Etienne-Cummings, "Giving machines humanlike eyes," *Spectrum, IEEE*, vol. 52, no. 12, pp. 44–49, 2015.

[296] S. Chhaniyara, P. Bunnun, L. D. Seneviratne, and K. Althoefer, "Optical Flow Algorithm for Velocity Estimation of Ground Vehicles: A Feasibility Study," *Int. J. Smart Sens. Intell. Syst.*, vol. 1, no. 1, pp. 246–268, 2008.

[297] N. Nourani-Vatani, J. Roberts, and M. V. Srinivasan, "Practical visual odometry for car-like vehicles," in *Proc. IEEE Int. Conf. Robot. Autom.*, pp. 3551–3557, 2009.

[298] M. Killpack, T. Deyle, C. Anderson, and C. C. Kemp, "Visual odometry and control for an omnidirectional mobile robot with a downward-facing camera," in *Proc. IEEE Int. Conf. Intell. Robot. Syst.*, pp. 139–146, 2010.

[299] J. Zhang, S. Singh, and G. Kantor, "Robust Monocular Visual Odometry for a Ground Vehicle in Undulating Terrain," in *Proc. Int. Conf. Field Serv. Robot.*, 2012.

[300] I. Nagai and K. Watanabe, "Path Tracking by a Mobile Robot Equipped with Only a Downward Facing Camera," pp. 6053–6058, 2015.

[301] S. Kim and S. Lee, "Robust velocity estimation of an omnidirectional mobile robot using a polygonal array of optical mice," *Int. J. Contr. Autom. Syst.*, vol. 6, no. 5, pp. 713–721, 2008.

[302] M. Dille, B. Grocholsky, and S. Singh, "Outdoor downward-facing optical flow odometry with commodity sensors," in *Proc. Int. Conf. Field Serv. Robot.*, pp. 183–193, 2009.

[303] R. Ross, J. Devlin, and S. Wang, "Toward refocused optical mouse sensors for outdoor optical flow odometry," *IEEE Sensors J.*, vol. 12, no. 6, pp. 1925–1932, 2012.

[304] I. Nagai, K. Watanabe, K. Nagatani, and K. Yoshida, "Noncontact position estimation device with optical sensor and laser sources for mobile robots traversing slippery terrains," in *Proc. IEEE Int. Conf. Intell. Robot. Syst.*, pp. 3422–3427, 2010.

[305] D.-H. Yi, T.-J. Lee, and D.-I. Cho, "Afocal Optical Flow Sensor for Reducing Vertical Height Sensitivity in Indoor Robot Localization and Navigation," *Sensors*, vol. 15, no. 5, pp. 11208–11221, 2015.

[306] S. Mafrica, A. Servel, and F. Ruffier, "Optic-flow based car-like robot operating in a 5-decade light level range," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 5568–5575, 2016.

[307] A. W. Snyder, "Acuity of compound eyes: Physical limitations and design," *Journal of comparative physiology*, vol. 116, no. 2, pp. 161–182, 1977.

[308] A. Manecy, N. Marchand, F. Ruffier, and S. Viollet, "X4-MaG : A Low-Cost Open-Source Micro-Quadrotor and its Linux-Based Controller," *Int. J. Micro Air Veh.*, vol. 7, no. 2, 2015.

[309] A. Martinelli, "Vision and imu data fusion: Closed-form solutions for attitude, speed, absolute scale, and bias determination," *IEEE Transactions on Robotics*, vol. 28, no. 1, pp. 44–60, 2012.

[310] Z. Wu, Z. Sun, W. Zhang, and Q. Chen, "Attitude and gyro bias estimation by the rotation of an inertial measurement unit," *Measurement Science and Technology*, vol. 26, no. 12, p. 125102, 2015.

[311] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? the kitti vision benchmark suite," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pp. 3354–3361, 2012.

[312] T. Zhang, X. Liu, K. Kuhnlenz, and M. Buss, "Visual odometry for the autonomous city explorer," in *Proc. IEEE Int. Conf. Intell. Robot. Syst.*, pp. 3513–3518, 2009.

[313] K. S. Chong and L. Kleeman, "Accurate odometry and error modelling for a mobile robot," in *Robotics and Automation, 1997. Proceedings., 1997 IEEE International Conference on*, vol. 4, pp. 2783–2788 vol.4, 1997.

Résumé de la thèse en français

Introduction

Grâce aux progrès réalisés dans les domaines de la robotique et des systèmes de transport intelligents (ITS), les véhicules autonomes du futur sont en train de devenir une réalité. Les véhicules autonomes devront se comporter en toute sécurité en présence d'autres véhicules, mais aussi de piétions et d'autres objets fixes ou en mouvement. Ainsi, une des choses les plus importantes est de percevoir efficacement à la fois leur mouvement et l'environnement autour d'eux. Le nombre et les types de capteurs embarqués sur les voitures donc se multiplient pour pouvoir fournir des informations redondantes et complémentaires et, éventuellement, d'atteindre des niveaux plus élevés d'autonomie et de sécurité.

Des solutions technologiques efficaces pourraient être trouvées en regardant la nature, qui suggère souvent des solutions simples mais intelligentes, permettant aux animaux de traiter seulement les informations nécessaires pour des tâches bien définies. Les insectes, par exemple, peuvent effectuer des maœuvres complexes et rapides et éviter les obstacles en utilisant uniquement des informations visuelles "bas niveau" et peu de ressources de calcul.

Dans cette thèse, nous avons d'abord étudié comment des capteurs visuels bio-inspirés, appelés Local Motion Sensors (LMS), qui mesurent le flux optique en 1-D en utilisant seulement quelques pixels sur la base du système visuel de la mouche, pourraient être utilisés pour améliorer les manœuvres de stationnement automatiques. A ce propos, nous avons développé une méthode de calcul bas-coût pour la détection et le suivi d'une place de parking en temps réel en utilisant seulement des mesures de flux optique en 1-D autour du véhicule, ainsi que la vitesse et l'angle de braquage du véhicule. Des simulations de parking très simplifié en 2-D ont d'abord été réalisées en utilisant le logiciel Matlab / Simulink, puis quelques expériences préliminaires ont été effectuées sur un véhicule équipé de deux capteurs composés de 6 pixels.

L'un des principaux défis pour les capteurs visuels est de fonctionner correctement dans une large gamme de conditions d'éclairage. Nous avons donc travaillé sur une nouvelle rétine de silicium auto-adaptative bio-inspirée conçue et développée par notre laboratoire en collaboration avec le Centre de Physique des Particules de Marseille (CPPM). Nous avons testé avec succès cette rétine, en montrant que le nouveau pixel, appelé M^{2}APIX (pour "Michaelis-Menten Auto-Adaptive Pixel"), est capable de s'auto-adaptater dans une gamme de 7 décades et de répondre de manière appropriée à des changements de luminosité rapides jusqu'à ± 3 décades, tout en conservant une sensibilité aux contrastes du 2%.

Enfin, nous avons développé et testé un nouveau capteur de flux optique basé sur cette rétine auto-adaptative et sur une nouvelle méthode robuste pour le calcul du flux optique, qui fournit plusieurs avantages par rapport à des capteurs développés précédemment, tel que sa robustesse aux variations de niveaux de lumière, textures et vibrations que l'on retrouve en milieu routier. Pour tester les performances de ce nouveau capteur et montrer comment il peut être utilisé pour des applications robotiques et automobiles tels que l'odométrie visuelle, nous avons construit un robot de type voiture, appelé BioCarBot, qui estime sa vitesse et son angle de braquage au moyen d'un filtre de Kalman étendu (EKF), en utilisant uniquement les mesures de flux optique délivrées par deux capteurs de ce type regardant vers le sol. Des expériences à l'intérieur et à l'extérieur ont été réalisées avec succès dans une gamme de luminosité de 7 décades et en utilisant différentes textures, montrant des perspectives prometteuses de ces capteurs pour des applications du type odométrie.

Capteurs visuels bio-inspirés pour des manœvres de parking automatiques

Depuis les travaux pionniers de Paromtchik et al. [251] et le premier système de stationnement intelligent [252], une large gamme de méthodes de parking (semi-)automatique ont été développées au fil des années en utilisant toutes sortes de capteurs, des capteurs à ultrasons [253–255] et caméras [256–258], aux systèmes à lidars [13, 259] et radars [15, 260].

D'une part, les systèmes de stationnement automatique de haute performance exigent des capteurs coûteux, tels que les lidars [13, 259] et les radars [15, 260], et de grandes ressources de calcul pour reconstituer l'environnement en 3-D en temps réel pendant les manoeuvres de stationnement.

D'autre part, la plupart des systèmes de stationnement (semi-)automatique disponibles sur le marché impliquent l'utilisation de capteurs à ultrasons et caméras de recul [8–10, 255, 257] pour détecter une place de parking libre pendant que le conducteur conduit le long d'une rangée de véhicules garés. Une fois la place détectée et validée par le conducteur, le véhicule effectue la manœvre de manière (semi-)autonome en contrôlant le braquage sur la base de la trajectoire estimée par l'odomètre.

Nous avons donc étudié comment les capteurs de flux optique bio-inspirés développés par notre laboratoire (LMS) [216, 242] pourraient être utilisé pour améliorer les manoeuvres de stationnement automatique. Grâce à leur large gamme de luminosité et leur sensibilité aux petits contrastes [242, 261], les LMSs comprenant seulement quelques pixels auto-adaptatifs semblent constituer une alternative prometteuse aux caméras standards dans des situations où les conditions d'éclairage sont très variables et les motifs visuels créés par les carrosseries des véhicules, par exemple, montrent de petits contrastes. En outre, un banc de LMSs pourrait présenter les avantages suivants par rapport aux capteurs à ultrasons et caméras: (i) des réponses plus rapides (jusqu'à 200 Hz); (ii) un large champ visuel (par exemple 180^o); (iii) une gamme de distance supérieure aux ultrasons; (iv) le coût de calcul plus faible des caméras.

Dans ce premier article, nous avons présenté les premiers résultats vers le développement d'une méthode de calcul à bas coût pour l'exécution de manoeuvres de stationnement automatiques, selon laquelle

une place de stationnement peut être détecté et suivi en temps réel sur une configuration du capteur de mouvement visuel effectuant des mesures en 1-D autour du véhicule. L'algorithme utilisé à ce but était composé de 5 étapes de traitement (Fig. I).

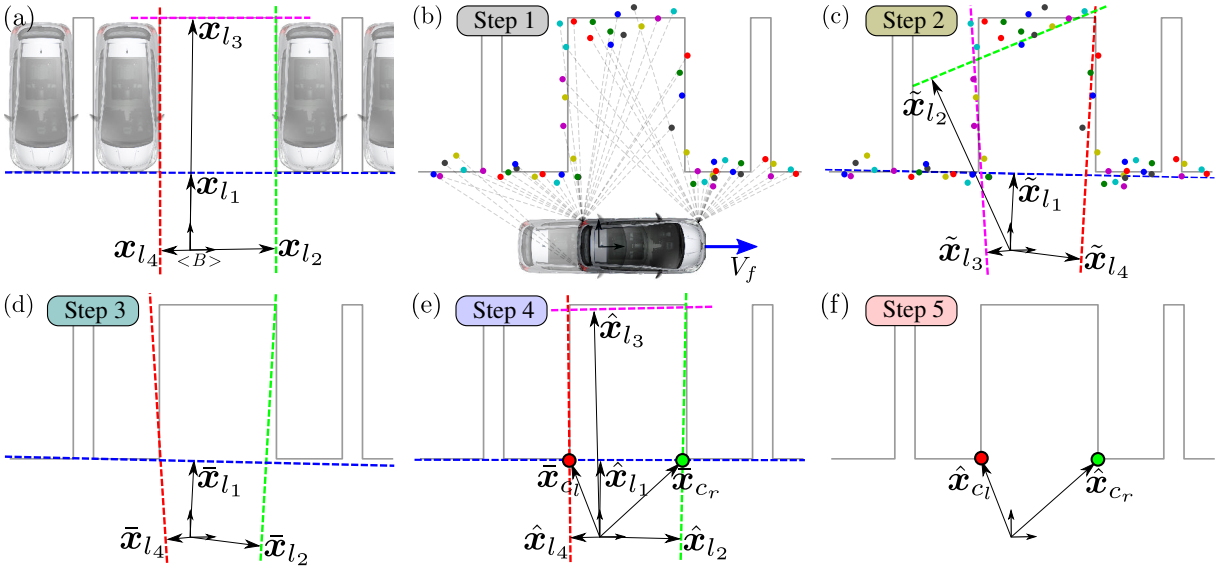

Figure I: Exemple de résultats obtenus à chaque étape de l'algorithme pour la détection et de suivi d'une place de parking à un intervalle d'échantillonnage au cours de la simulation. (a) La place de parking simplifiée en 2-D avec les vecteurs de position perpendiculaires aux 4 lignes droites définissant la place. (b) Les points détectés sur les axes des pixels, dont les positions ont été calculées en utilisant les mesures de flux optique en 1-D délivrées par les LMSs placés aux 4 coins du véhicule (étape 1). Un bruit blanc a été ajouté aux positions de ces points afin d'ajouter un bruit non linéaire aux mesures du flux optique. (c) Les vecteurs de position des lignes obtenues par RANSAC dans le nuage de points (étape 2). (d) Les vecteurs de position des lignes classifiés par les NBCs (étape 3). (e) Les vecteurs de position des lignes estimés par les EKFs (étape 4). Une première approximation de la position des coins de stationnement est donnée par les points d'intersection entre chaque paire de droites perpendiculaires. (f) Les vecteurs des positions des coins de la place de parking estimés par les EKFs (étape 5).

D'abord, l'algorithme calcule les positions des points détectés par les LMSs en utilisant seulement les mesures de flux optique en 1-D autour du véhicule, ainsi que la vitesse et l'angle de braquage du véhicule (étape 1), puis il cherche des lignes droites dans les nuages de points en utilisant la méthode de RANSAC [262] (étape 2). Les lignes qui ont été trouvées sont ensuite classifiées par des Naive Bayes Classifiers (NBCs) pour assurer une cohérence avec la géométrie simplifiée de la place de parking simulée (étape 3). Enfin, les vecteurs perpendiculaires aux lignes de la place de parking et les positions des points d'intersections entre ces lignes (coins de la place) sont estimés à l'aide des filtres de Kalman étendus (EKFs) (étapes 4 and 5). Une stratégie de contrôle non-linéaire peut-être finalement appliquée sur la base de l'estimation de la position du véhicule par rapport aux lignes et coins de la place de parking, mais cela n'a pas été le point principal de ce travail.

Des résultats obtenus en simulation dans un parking simplifié en 2-D, mis en oeuvre dans Matlab /

Simulink en boucle fermée, ont été d'abord présenté. Puis, nous avons montré des résultats expérimentaux préliminaires correspondant à la première étape de l'algorithme dans le cas d'un véhicule (Peugeot 3008) équipé de 2 capteurs à 6 pixels attachés sur le coin droit du pare-choc avant (Fig. II).

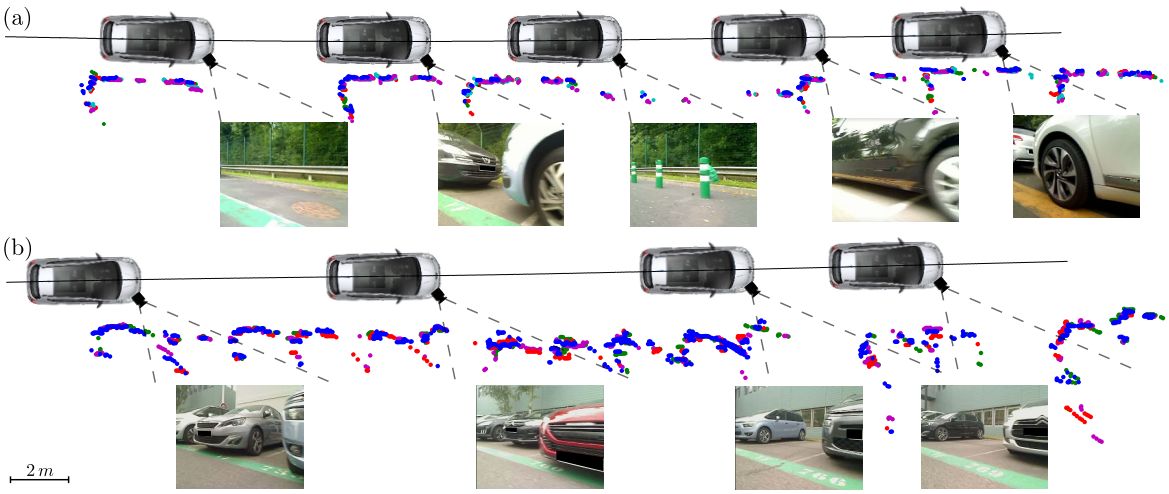

Figure II: Exemple de trajectoire du véhicule, ainsi que les positions des points en 2-D calculées en utilisant le flux optique mesuré par l'un des deux LMSs à 6 pixels attachés au pare-chocs du véhicule ($\Delta\varphi \approx 3.8^o$ [242]), lorsque le véhicule se déplaçait à une vitesse entre 3 et 9 $\frac{m}{s}$ le long des véhicules garés (a) en créneau et (b) en bataille. Les images dans la partie inférieure de chaque sous-figure sont les images prises par la webcam lorsque le véhicule était dans les positions présentées dans la partie supérieure, et les lignes en pointillés donnent une indication approximative du champ visuel de la webcam ($\sim 45^o$).

La fréquence de rafraîchissement du flux optique produit par les véhicules stationnés était entre 30 et 65 Hz, ce qui est environ 2-3 fois supérieures à celle des capteurs à ultrasons et caméras standards. Dans les deux cas de parking parallèle et perpendiculaire, il est souvent possible de reconnaître les formes des véhicules garés dans les nuages de points, même si peu de mesures sont parfois obtenues en raison, par exemple, des réflexions sur les carrosseries de voiture (voir, par exemple, la quatrième image dans la Fig. II (a)). Dans le cas de stationnement perpendiculaire, les mesures étaient plus bruyants que dans le cas parallèle, principalement en raison (i) des occlusions des côtés des voitures et (ii) des désalignements entre les voitures (voir, par exemple, la première et seconde image dans la Fig. II (b)).

Les résultats qualitatifs présentés ici valident la première étape de l'algorithme que nous avons développé. Bien qu'il n'a pas été possible, dans ces premiers tests, d'appliquer les autres étapes de l'algorithme parce qu'il y avait trop peu de points dans chaque intervalle d'échantillonnage, il semble que, à l'aide de LMSs avec un plus grand nombre de pixels (ou plusieurs LMSs à 6 pixels), nous aurons probablement assez de points pour être en mesure d'appliquer toutes les étapes.

M^{2}APix: un nouveau pixel auto-adaptatif bio-inspiré

Au cours des dernières décennies, la recherche dans le domaine de la robotique a considérablement progressé, mais il existe encore très peu de robots basés sur la vision qui sont capables de se comporter de manière appropriée, indépendamment des changements d'éclairage (voir Fig. 5 dans [242], par exemple), tels que ceux qui se produisent à l'extérieur. Une des raisons de ce manque est la difficulté de concevoir des pixels qui combinent haute sensibilité au contraste avec une grande gamme luminosité.

Une grande variété d'imageurs CMOS Wide-Dynamic-Range (WDR) a été proposé tout au long des années [156], en essayant d'élargir la gamme de luminosité de fonctionnement autant que le spectre visible, tout en conservant une sensibilité aux petits changements pour chaque niveau de luminosité moyenne. Bien que les imageurs WDR puissent capturer des images dans une gamme de luminosité allant jusqu'à 7 décades, ils fournissent des sensibilités au contraste qui dépendent de l'éclairage moyen. Des applications visuelles, telles que les applications évènementielles et biomédicales, nécessitent souvent une sensibilité élevée et constante dans une large gamme de luminosité, afin de détecter de petites variations d'intensité temporelles et/ou spatiales dans plusieurs conditions d'éclairage [169, 170, 265, 266]. Une solution possible à ce problème peut être trouvée en examinant la réponse auto-adaptative des photorécepteurs chez les humains et les animaux.

Pour autant que nous savons, aucune rétine artificielle n'a jamais été dotée jusqu'à présent de pixels avec toutes les caractéristiques suivantes à la fois : (i) une auto-adaptation à la luminosité moyenne locale sur une gamme aussi large que le spectre visible; (ii) une bonne sensibilité aux variations lumineuses, soit les contrastes, à tout niveau de luminosité moyenne dans la plage de fonctionnement; (iii) une réponse fiable, même en présence de fortes variations brusques de luminosité (à savoir, sans provoquer des saturations des transistors VLSI).

Dans cet article, nous avons présenté: (i) une nouvelle rétine de silicium analogique ayant des pixels auto-adaptatifs qui obéissent à la loi de Michaelis-Menten, à savoir $V = V_m \frac{I^n}{I^n+\sigma^n}$; (ii) une méthode de caractérisation de rétines de silicium, ce qui permet d'évaluer avec précision la réponse des pixels aux changements lumineux transitoires dans une gamme de ± 3 décades, ainsi que des changements de lumière ambiante dans une gamme de 7 décades.

La rétine artificielle présentée ici est composée d'un circuit CMOS de taille $2 \times 2\,mm$, comprenant quatre lignes de 6 pixels auto-adaptatifs chacune, 2 lignes de type M^{2}APix et 2 lignes de type Delbrück (Fig. III). La rétine de silicium présente aussi une interface numérique pour envoyer les données des pixels en série (jusqu'à 1 MHz) directement à un microprocesseur ou d'un microcontrôleur externe. La constante de temps d'adaptation du pixel M^{2}APix peut être modifiée au moyen d'un condensateur externe, offrant une plus grande souplesse pour répondre éventuellement aux besoins de l'application en terme de bande passante souhaitée.

Le nouveau pixel, appelé M^{2}APix, qui signifie Michaelis-Menten Auto-Adaptive Pixel, montre en termes qualitatifs le même processus d'adaptation que celui observé dans la rétine des humains et des animaux (Fig. IV). En particulier, il peut s'auto-adaptater dans une gamme de 7 décades et réagit de façon

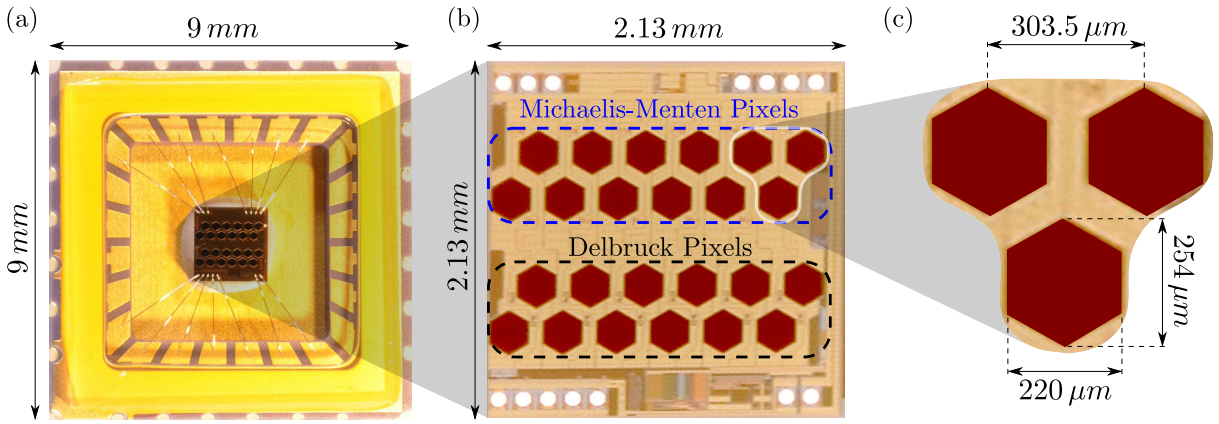

Figure III: (a) Photographie de la rétine artificielle dans son package de $9 \times 9\,mm$. (b) Vue agrandie de la rétine composé de 12 pixels Michaelis-Menten (M^2APix) présentés dans cette étude, et 12 pixels supplémentaires de type Delbrück. (c) Vue agrandie de 3 M^2APix avec les dimensions de la photodiode et les distances inter-récepteurs.

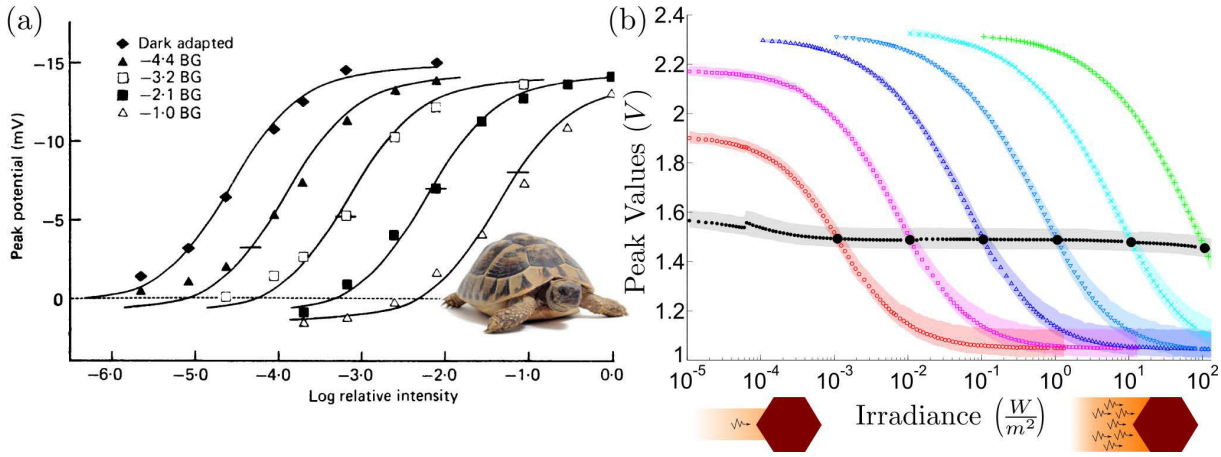

Figure IV: (a) Les courbes en S correspondant aux réponses enregistrées dans un cône rouge de la tortue, lorsque des changements de luminosité (flashs) sont produits une fois que les photorécepteurs se sont adaptés à l'obscurité ou à la lumière. Les courbes continues ont été tirées à partir du modèle décrit par la fonction $V = V_m \frac{I}{I+\sigma}$. Adapté de [47]. (b) Les courbes en S et la réponse statique de 12 pixels M^2APix en fonction de l'irradiance de la LED dans une gamme de 7 décades. Chaque couleur se réfère à une valeur d'irradiance initiale différente (grands cercles noirs), et les points correspondent à la moyenne des valeurs de pic atteint par les 12 pixels en réponse à un changement d'irradiance en échelon. La réponse statique (points noirs) a été obtenue en maintenant pendant longtemps la même irradiance. Les zones ombrées ont été obtenues en traçant les valeurs minimales et maximales des tensions de sortie des pixels.

appropriée à des changements brusques jusqu'à ± 3 décades, sans provoquer de saturation des transistors en VLSI. Grâce aux propriétés intrinsèques de l'équation de Michaelis-Menten, la sortie de pixel reste toujours dans une plage de tension limitée et constante. La plage du convertisseur analogique-numérique (ADC) a donc été ajusté de manière à obtenir une tension correspondante au bit le moins significatif (LSB) de $2.35\,mV$, et une résolution effective d'environ 9 bits. Les résultats présentés ici montrent que les pixels M^{2}APix produisent une réponse au contraste quasi-linéaire une fois qu'il se sont adaptés à la luminosité moyenne. Différemment de ce qui se passe dans ses homologues biologiques, ni la sensibilité aux changements de lumière, ni la réponse au contraste de M^{2}APix dépendent de la luminosité moyenne (à savoir, les conditions d'éclairage ambiant). Enfin, une comparaison complète entre le pixel M^{2}APix et le pixel auto-adaptatif de Delbrück a été effectuée.

Un nouveau capteur de flux optique bio-inspiré et son application à l'odométrie visuelle

Le flux optique (OF) a été étudié par de nombreux auteurs au cours des dernières décennies [100, 109, 282,283] : divers capteurs [177,192,209] et algorithmes [98,99,111] ont été développés et utilisés pour des applications robotiques telles que la navigation autonome [233, 284], le contrôle de vitesse [242, 285], la localisation et la cartographie simultanée (SLAM) [21, 22], et l'odométrie visuelle [286–289].

Des méthodes d'odométrie visuelle pour les robots mobiles et les automobiles ont été récemment mis au point à l'aide de caméras standards [296–300] et des souris optiques [301–303], ainsi que de capteurs customs [304, 305], orientés vers le bas, puisque les motifs visuels et les conditions de lumière rencontrées dans ce contexte sont relativement uniformes. Dans la plupart des cas, les auteurs ont profité de la contrainte non-holonome de type voiture pour estimer la position et l'orientation du véhicule en utilisant des méthodes simples impliquant des coûts de calcul faibles.

Cependant, des solutions basées sur les caméras standards [297, 300] ne parviennent pas encore à faire face à des conditions d'éclairage en haute gamme dynamique, et ont toujours le problème d'une cadence faible et un coût de calcul du traitement d'image élevé : seulement une petite gamme de vitesse faible peut souvent être obtenue en utilisant cette approche.

Les solutions basées sur des souris optiques [301, 303] sont certainement très économiques et offrent des mesures à haute fréquence, mais leur principal inconvénient est qu'ils doivent opérer très près du sol pour pouvoir fonctionner correctement, et ne conviennent donc pas pour une utilisation dans des environnements avec un terrain accidenté. En outre, ces capteurs sont généralement très sensibles aux conditions d'éclairage et comme les caméras standards, fournissent des mesures dans une plage de vitesse assez petite.

Des solutions basées sur des capteurs customs éclairés par LED ou laser [304,305] ont été développées afin de réduire la sensibilité à la hauteur et d'améliorer les performances tout en voyageant sur des terrains de types différents, mais aucun test n'a été effectué par ces auteurs dans différentes conditions d'éclairage et dans des conditions opérationnelles du robot.

Dans cet article, nous avons présenté:

- un nouveau capteur bio-inspiré donnant des mesures de flux optique qui sont robustes aux conditions d'éclairage dans une haute gamme dynamique, et qui ne dépendent presque pas des motifs visuels rencontrés;
- une application à l'odométrie visuelle, dans laquelle ces nouveaux capteurs ont été installés sur un robot bas-coût de type voiture, appelé BioCarBot (Fig. V), et testés dans une gamme de niveau de lumière de 7 décades.

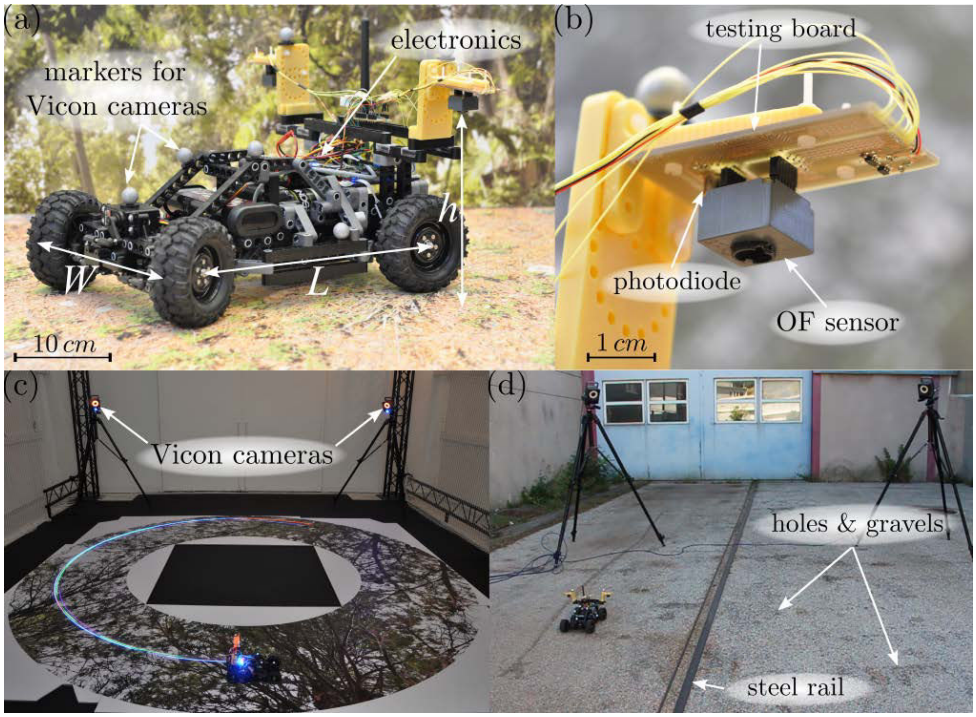

Figure V: (a) Photographie du robot bas-coût de de type voiture appelé BioCarBot. (b) Photographie de l'un des 2 capteurs de flux optique utilisés sur le robot. (c) Photographie de l'environnement de test intérieur équipé de caméras Vicon. (d) Photographie de l'environnement de test extérieur équipé de caméras Vicon. Le terrain se composait principalement d'asphalte, avec des trous, du gravier et une raille en métal.

Un filtre de Kalman étendu (EKF) a été utilisé pour estimer la vitesse du robot et son angle de braquage en utilisant uniquement des mesures de flux optique délivrées par deux capteurs tournées vers le bas. La position et l'orientation absolues du robot ont ensuite été estimées en temps réel en utilisant ces estimations du EKF et le modèle d'Ackerman [264]. Des expériences à l'intérieur et à l'extérieur ont été réalisées, dans lequel le robot a été commandé en boucle fermée sur plusieurs types de sols et de trajectoires.

Les résultats présentés ici montrent que nos nouveaux capteurs de flux optique sont robustes aux changements de lumière dans une plage de 7 décades (de environ $10^{-10}\,A$ à $10^{-3}\,A$ du courant de photodiode), y compris des changements brusques jusqu'à 2 décades en $0,5\,s$ (Fig. VI). Bien que des

servos et des moteurs à bas coût et à faible résolution, ainsi qu'une identification et calibration du modèle simplifiées, ont été utilisés, le robot a pu estimer sa vitesse et son angle de braquage de façon précise. La position et l'orientation du robot ont été estimées à l'intérieur et à l'extérieur alors que le robot se déplaçait dans des environnements non structurés (sur un sol constitué par l'asphalte, le gravier, le sable et les feuilles, y compris des ombres et des trous), que ça soit pendant la journée ou la nuit. En outre, la méthode d'odométrie visuelle s'est montrée être robuste à des vibrations susceptibles de modifier la hauteur locale des capteurs jusqu'au 6% (à savoir, $\pm 10\,mm$ sur $175\,mm$).

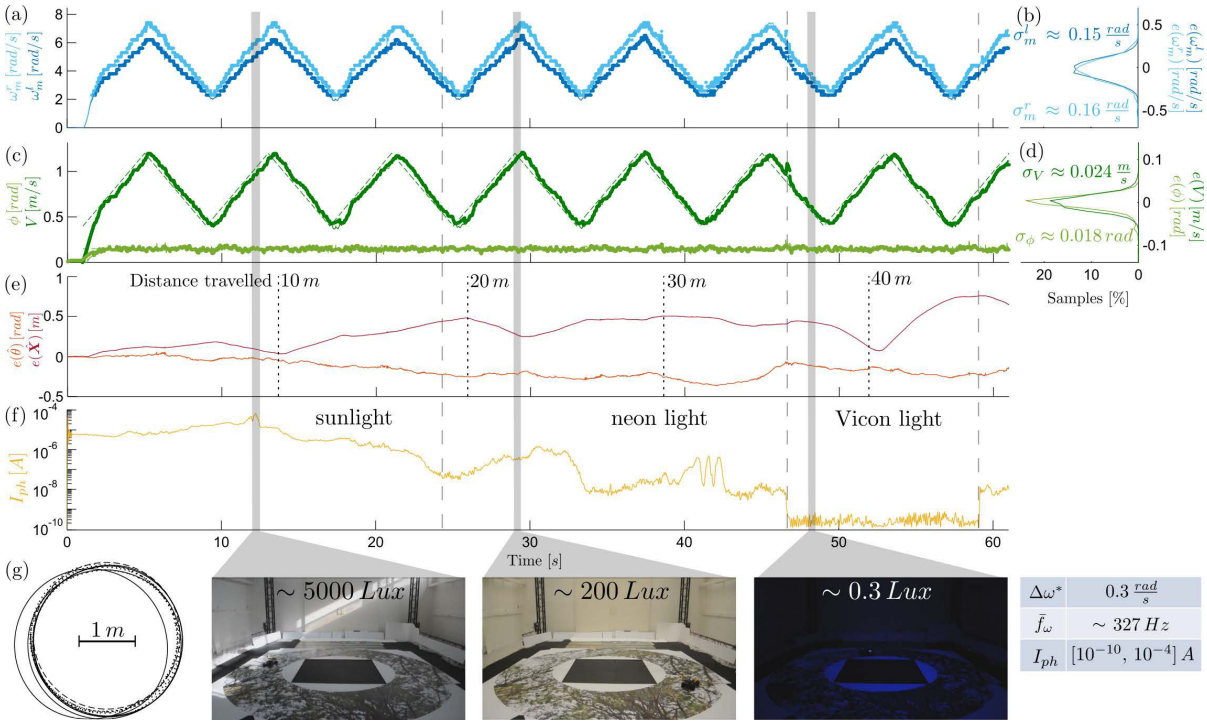

Figure VI: Mesures de flux optique et résultats d'odométrie robustes à des changements de lumière dans une gamme d'environ 7 décades, obtenus lorsque le robot a été conduit sur une trajectoire circulaire à une vitesse allant de 0.3 à 1.3 $\frac{m}{s}$.. (a) Valeurs médianes ω^l_m, ω^r_m (points) et valeurs de vérité-terrain $\omega^l_{m_{truth}}, \omega^r_{m_{truth}}$ (traits pleins) du flux optique produit sur côté gauche (bleu foncé) et droit (bleu clair). (b) Distribution des erreurs entre les valeurs médianes et les valeurs de vérité-terrain montrées en Fig. 4.6 (a), soit $\omega^l_{m_{truth}} - \omega^l_m$ (ligne bleu foncé) et $\omega^r_{m_{truth}} - \omega^r_m$ (ligne bleu clair). (c) Estimations de l'EKF $\hat{V}, \hat{\phi}$ (points), valeurs de vérité-terrain V_{truth}, ϕ_{truth} (lignes continues) et valeurs de référence V^*, ϕ^* (lignes pointillées) de la vitesse (vert foncé) et de l'angle de braquage (vert clair) du robot. (d) Distribution des erreurs entre les estimations de l'EKF et les valeurs de vérité-terrain montrées en Fig. 4.6(c), soit $V_{truth} - \hat{V}$ (ligne vert foncé) et $\phi_{truth} - \hat{\phi}$ (ligne vert clair). (e) Erreurs de position et orientation du robot estimées en utilisant $\hat{V}, \hat{\phi}$, i.e. $||\boldsymbol{X} - \hat{\boldsymbol{X}}||$ (ligne rouge foncé) et $\theta - \hat{\theta}$ (ligne rouge clair). (f) Réponse photo-courant de la photodiode placée à côté du LMS sur le côté gauche. Ces valeurs ont été calculés en prenant $I_{Ph} = I_{dark}(e^{8.8V_{out}} - 1)$, où $I_{dark}(\approx 0.1\,nA)$ est le courant d'obscurité et V_{out} est la tension de sortie de l'amplificateur. (g) Valeurs de référence $\boldsymbol{X}^*$ (ligne pointillée), estimations $\hat{\boldsymbol{X}}$ (ligne pointillée) et vérité-terrain $\boldsymbol{X}$ (ligne continue) de la trajectoire du robot en boucle fermée. Les images au bas de la figure montrent, de gauche à droite, l'environnement de test sous la lumière du soleil entrant, la lumière des néons et la lumière des LEDs des caméras Vicon. Les valeurs de lumière en Lux ont été estimées en utilisant la courbe courant-éclairement donné dans la fiche technique de la photodiode.

Conclusion

Dans cette thèse, nous avons voulu proposer des nouvelles solutions technologiques bio-inspirées qui pourraient être utilisées dans les systèmes d'aide à la conduite (ADAS), et en particulier sur les véhicules autonomes du futur, pour lesquels redondance et complémentarité de l'information perçue sera nécessaire pour satisfaire les contraintes d'autonomie et de sécurité.

Comme indiqué dans le premier chapitre de ce manuscrit, les systèmes ADAS existants offrent déjà un certain nombre de degrés d'assistance au conducteur, mais, dans leur forme actuelle, ne sont pas encore en mesure de fournir des expériences de conduite autonome qui sont complètes et à prix compétitif. Certains prototypes des premiers véhicules autonomes ont déjà été présentés, mais ces solutions sont encore trop complexes et coûteuses, et exigent souvent trop grandes ressources de calcul pour être embarquées sur des voitures de série.

Nous croyons que des solutions technologiques efficaces pourraient être trouvées en s'inspirant de la nature, qui suggère souvent des solutions simples, permettant aux animaux de traiter seulement les informations nécessaires pour des tâches bien définies. Les capteurs de mouvement local (LMS), conçus et développés dans notre laboratoire utilisent très peu de pixels et de ressources de calcul, s'inspirant ainsi du système visuel de la mouche.

Bien que les LMSs proposés jusqu'à présent ont prouvé leur efficacité dans la mesure flux optique en 1-D dans certaines situations / environnements, et ont été utilisé avec succès sur des robots mobiles et aériennes pour effectuer de l'évitement d'obstacles et du contrôle de vitesse, ils nécessitent encore quelques améliorations pour fonctionner correctement sur des vrai véhicules dans des environnements routiers. Les principales améliorations nécessaires étaient liées à (i) leur robustesse par rapport aux conditions d'éclairage, aux textures visuelles et aux vibrations, et (ii) leur capacité à fournir des mesures à un taux de rafraîchissement élevé et quasi-constante, avec une résolution fixe et constante.

Deux aspects ont été traités donc dans cette thèse: un aspect technologique, impliquant l'adaptation des LMSs à l'environnement routier et aux conditions de fonctionnement du véhicule, et un aspect applicatif, impliquant l'utilisation efficace des LMSs dans les systèmes ADAS.

- Dans la première partie, nous avons montré que, en utilisant une méthode "temps de passage" robuste pour le calcul du flux optique, basée sur la corrélation des signaux des pixels voisins, les LMSs peuvent éventuellement fournir des mesures à haute fréquence lorsqu'ils sont utilisés sur un véhicule dans des contextes de parking, même en présence de vibrations et peu de contrastes sur les carrosseries des voitures. Nous avons également montré que, dans les situations où les LMSs ont livré des mesures relativement précises, notamment pour le parking parallèle, il était possible de calculer des nuages de points en 2-D à la même haute fréquence, dans lesquels on peut y reconnaitre les silhouettes des voitures garées. Dans le cas où plusieurs LMSs pourraient être installés tout autour du véhicule, les lignes correspondant aux véhicules garés pourraient être trouvées en temps réel et la trajectoire du véhicule peut être estimée de manière simple et robuste pendant toute la manœuvre de stationnement, comme montré en simulation. Cependant, la réponse de ces

capteurs n'était toujours pas satisfaisante dans des conditions d'éclairage extrêmes, surtout à cause des signaux délivrés par les pixels dans telles conditions.

- Dans la deuxième partie, nous avons présenté une nouvelle rétine de silicium auto-adaptative composée de nouveaux pixels, appelés M^{2}APix, qui reproduisent les réponses auto-adaptatives enregistrées dans la rétine de la plupart des animaux et même des humains. Comme le fait de tester des pixels auto-adaptatifs et de les comparer avec d'autres ne nous parait pas une opération facile et bien définie, nous avons également proposé une nouvelle méthode standard pour caractériser précisément les réponses des pixels aux changements lumineux allant jusqu'à ± 3 décades dans une gamme de 7 décades de luminosité moyenne, en utilisant une seule source de lumière, appelée Lighting Box. Le nouveau M^{2}APix pixel a montré des résultats très prometteurs au cours des tests de caractérisation, présentant notamment aucune déviation de sa réponse en présence de changements lumineux grands et brusques, ce qui est un problème commun des puces VLSI. Toutefois, d'autres tests dans des conditions réelles de fonctionnement devraient être effectués pour prouver leur efficacité pour les mesures de flux optique.

- Dans la troisième et dernière partie, un nouveau capteur de flux optique robuste, basé sur la nouvelle rétine auto-adaptative et la nouvelle méthode de calcul de flux optique basée sur corrélation, a donc été construit, après l'amélioration de la de l'algorithme pour fournir une résolution constante et une mise en œuvre en électronique embarquée. Ce nouveau LMS a été testé avec succès sur un robot de type voiture à bas coût, appelé BioCarBot, pour estimer sa vitesse et son angle de braquage, et donc estimer sa trajectoire (odométrie), dans des différentes conditions de fonctionnement (niveaux de lumière, textures, vibrations, etc.) semblables à celles qui se produisent sur les véhicules routiers. Une estimation fiable et précise de la vitesse du véhicule est, en fait, crucial car elle est utilisée par la plupart des systèmes ADAS : l'efficacité des systèmes ABS et ESC, par exemple, dépend fortement de cette estimation, mais aussi celle de tout système de contrôle de la trajectoire du véhicule sur la base de l'odométrie (par exemple dans les systèmes d'aide au stationnement). Avec les expériences réalisées sur le robot BioCarBot nous avons montré que les LMSs développées ont pu faire face aux principaux problèmes rencontrés avec des versions précédentes de LMSs, comme l'amélioration de leur réponse à l'extérieur dans des environnements non structurés, et pourraient donc fournir une solution très prometteuse pour estimer la vitesse précisément, ainsi que pour des applications basées sur l'odométrie.

BIO-INSPIRED VISUAL SENSORS FOR ROBOTIC AND AUTOMOTIVE APPLICATIONS

ABSTRACT: Thanks to the advances in the fields of robotics and intelligent transportation systems (ITS), the autonomous vehicles of the future are gradually becoming a reality. As autonomous vehicles will have to behave safely in presence of other vehicles, pedestrians and other fixed and moving objects, one of the most important things they need to do is to effectively perceive both their motion and the environment around them. In this thesis, we first investigated how bio-inspired visual sensors, called Local Motion Sensors (LMSs), giving 1-D optic flow using a few pixels based on the findings on the fly's visual system, could be used to improve automatic parking maneuvers. For this purpose, we developed a low computational-cost method for detecting and tracking a parking spot in real time using only 1-D OF measurements around the vehicle together with the vehicle's longitudinal velocity and steering angle. Highly simplified 2-D parking simulations were first performed using Matlab/Simulink software, then some preliminary experiments were carried out using a vehicle equipped with two 6-pixel LMSs. As the main challenge for visual sensors is to correctly operate in high-dynamic-range lighting conditions, we also dealt here with a novel bio-inspired auto-adaptive silicon retina designed and developed by our laboratory in collaboration with the Center of Particle Physics of Marseille (CPPM). We successfully tested this silicon retina, showing that the novel pixel, called M^2APix, which stands for Michaelis-Menten Auto-Adaptive Pixel, can auto-adapt in a 7-decade range and respond appropriately to step changes up to ± 3 decades, while keeping sensitivity to contrasts as low as 2%. We subsequently developed and tested a novel optic flow sensor based on this auto-adaptive retina and a new robust method for computing the optic flow, which provides several advantages to previously developed optic flow sensors such as its robustness to light levels, textures and vibrations that can be found while operating on the road. To test the performances of this novel sensor and show how it can be used for robotic and automotive applications such as visual odometry, we constructed a car-like robot, called BioCarBot, which estimates its velocity and steering angle by means of an extended Kalman filter (EKF) using only the optic flow measurements delivered by two downward-facing sensors of this kind. Indoor and outdoor experiments were successfully carried out in a 7-decade light level range and using various textures, showing promising perspectives of these sensors for odometry-based applications.

KEYWORDS: Bio-Inspiration, Robotics, Vision, Optic Flow, Autonomous Vehicles, ADAS

CAPTEURS VISUELS BIO-INSPIRÉS POUR DES APPLICATIONS ROBOTIQUES ET AUTOMOBILES

RÉSUMÉ: Grâce aux progrès réalisés dans les domaines de la robotique et des systèmes de transport intelligents (ITS), les véhicules autonomes du futur sont en train de devenir une réalité. Comme les véhicules autonomes devront se comporter en toute sécurité en présence d'autres véhicules, mais aussi de piétions et d'autres objets fixes ou en mouvement, une des choses les plus importantes qu'ils doivent faire est de percevoir efficacement à la fois leur mouvement et l'environnement autour d'eux. Dans cette thèse, nous avons d'abord étudié comment des capteurs visuels bio-inspirés, appelés Local Motion Sensors (LMS), qui mesurent le flux optique en 1-D en utilisant seulement quelques pixels sur la base du système visuel de la mouche, pourraient être utilisés pour améliorer les manœuvres de stationnement automatiques. A ce propos, nous avons développé une méthode de calcul bas-coût pour la détection et le suivi d'une place de parking en temps réel en utilisant seulement des mesures de flux optique en 1-D autour du véhicule, ainsi que la vitesse longitudinale et l'angle de braquage du véhicule. Des simulations de parking très simplifié en 2-D ont d'abord été réalisées en utilisant le logiciel Matlab / Simulink, puis quelques expériences préliminaires ont été effectuées sur un véhicule équipé de deux LMSs composés de 6 pixels. Comme le principal défi pour les capteurs visuels est de fonctionner correctement dans des conditions d'éclairage dans une large gamme dynamique, nous avons ensuite travaillé sur une nouvelle rétine de silicium auto-adaptative bio-inspirée conçue et développée par notre laboratoire en collaboration avec le Centre de Physique des Particules de Marseille (CPPM). Nous avons testé avec succès cette rétine, en montrant que le nouveau pixel, appelé M^2APIX, qui signifie Michaelis-Menten Auto-Adaptive Pixel, est capable de s'auto-adapter dans une gamme de 7 décades et de répondre de manière appropriée à des changements de luminosité rapides jusqu'à ± 3 décades, tout en conservant une sensibilité aux contrastes aussi bas que 2%. Nous avons enfin développé et testé un nouveau capteur de flux optique basé sur cette rétine auto-adaptative et sur une nouvelle méthode robuste pour le calcul du flux optique, qui fournit plusieurs avantages par rapport à des capteurs développés précédemment, tel que sa robustesse aux variations de niveaux de lumière, textures et vibrations que l'on retrouve en milieu routier. Pour tester les performances de ce nouveau capteur et montrer comment il peut être utilisé pour des applications robotiques et automobiles tels que l'odométrie visuelle, nous avons construit un robot de type voiture, appelé BioCarBot, qui estime sa vitesse et son angle de braquage au moyen d'un filtre de Kalman étendu (EKF), en utilisant uniquement les mesures de flux optique délivrées par deux capteurs de ce type regardant vers le sol. Des expériences à l'intérieur et à l'extérieur ont été réalisées avec succès dans une gamme de luminosité de 7 décades et en utilisant différentes textures, en montrant des perspectives prometteuses de ces capteurs pour des applications basées sur odométrie.

MOTS CLÉS: Bio-Inspiration, Robotique, Vision, Flux Optique, Véhicules Autonomes, ADAS