Computer Vision, Reinforcement Learning and Skill Transfer in Robotic: Application to a Dynamic Task

presented by

Antonio Claudio DE SOUSA DOS SANTOS FILHO

Internship Report

Realized from 03 April 2023 to 28 July 2023

Company tutors: Sébastien LENGAGNE and Samuel BEAUSSANT
Academic tutor: Nancy EL RACHKIDY

July 2023

Polytech Clermont-Université Clermont Auvergne

Campus Universitaire des Cézeaux – 2, avenue Blaise-Pascal

63178 Aubière Cedex

Tél. (33) 04 73 40 75 00

https://polytech-clermont.fr/

Abstract

This report presents the application of Computer Vision and Reinforcement Learning techniques for ball detection and catching using an event based camera. The internship aimed to develop an efficient method for detecting ball trajectories and contributed to the transfer learning research of Samuel Beaussant's doctoral thesis. This report provides an introduction to Computer Vision, Reinforcement Learning, and skill transfer in robotics. It describes the organization where the internship took place, the Institut Pascal at the Université Clermont Auvergne. The methodology, internship development, collaboration with the PhD thesis, and results discussion are presented.

keywords: Computer Vision, Reinforcement Learning,Transfer Learning, event based camera

Résumé

Ce rapport présente l'application des techniques de vision par ordinateur et d'apprentissage par renforcement pour la détection et la capture de balles à l'aide d'une caméra événementielle. Le stage visait à développer une méthode efficace pour détecter les trajectoires de balles et contribuer à la recherche sur l'apprentissage par transfert de la thèse de doctorat de Mr. Samuel Beaussant. Ce rapport fournit une introduction à la vision par ordinateur, à l'apprentissage par renforcement et au transfert de compétences en robotique. Il décrit l'organisation où le stage a eu lieu, l'Institut Pascal de l'Université Clermont Auvergne. La méthodologie, le développement du stage, la collaboration avec la thèse de doctorat et la discussion des résultats sont présentés.

Mots Clés : Vision par ordinateur, Apprentissage par Renforcement, Transfert d'apprentissage, Caméra événementielle

Acknowledgement

Firstly, I would like to express my gratitude to my supervisor, Mr. Sébastien Lengagne, for providing me with the opportunity to undertake this internship in his team. His guidance, expertise, and encouragement have been essential to my professional and academic growth throughout this period. I would like to thank my academic tutor, Ms. Nancy El Rachkidy, for their constant support throughout this internship.

I would also like to thank Mr. Samuel Beaussant, the PhD student, for the opportunity and the knowledge he shared with me regarding his thesis.

I would like to extend my special thanks to Professor Céline Teulière, of the ComSee lab and her PhD student Mr. Antony N'Dri, for their support and guidance throughout this internship. Their expertise and insight have been invaluable in developing my knowledge and skills in the field of Computer Vision.

Finally, I would like to thank the PhD students present in the laboratory, Mr. Bastien Mourin, Mr. Jules Caposiena, and Mr. Victor Giraud. Their contributions, discussions, and knowledge exchange enriched my experience in the laboratory and expanded my understanding in their respective research areas.

Summary

List of Figures

Acronyms

DBSCAN Density-Based Spatial Clustering of Applications with Noise. 2, 18, 19

DoF Degrees of Freedom. 2, 3, 7, 23

MDP Markov Decision Process. 23

ML Machine Learning. 6, 11, 29

PPO Proximal Policy Optimisation. 24

RL Reinforcement Learning. 6, 7, 11–13, 24, 29

UNN Universal Notice Network. 2, 13, 14, 25

Nomenclature

cm Centimeters

ms Milliseconds

us Microseconds

CHAPTER 1

Introduction

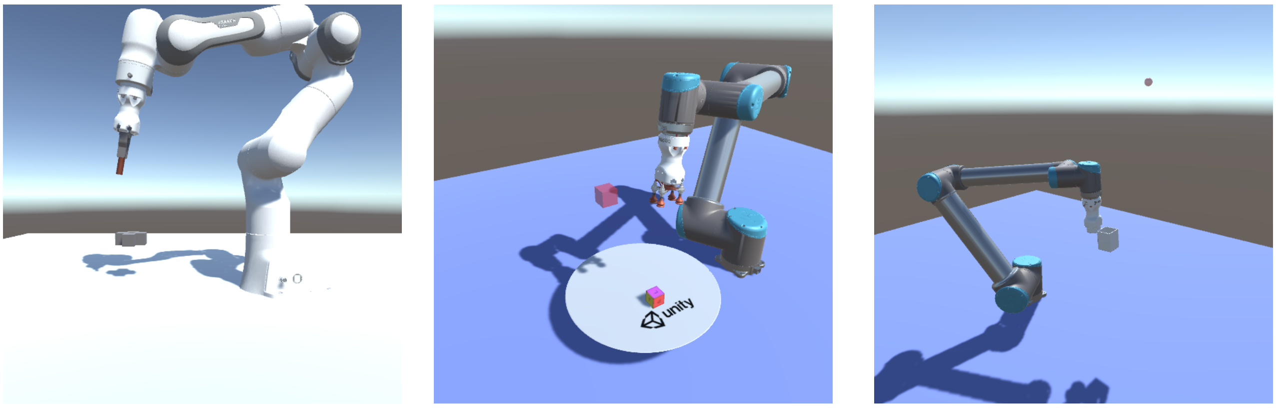
The field of Machine Learning (ML) has gained prominence in recent years. This notoriety can be attributed to the potential to enhance the autonomy and efficiency of computers in a wide variety of domains such as solving complex tasks shown in [1] or more recently the results obtained with ChatGPT. In this context, the current stage aims to explore and apply areas such as Reinforcement Learning (RL), Computer Vision and Transfer Learning, to develop dynamic tasks, which will serve as a contribution to Samuel Beaussant's doctoral thesis.

The internship is focused on contributing to Samuel Beaussant's doctoral thesis, specifically in the validation of Transfer Learning in the dynamic tasks, in our case in the tasks Ball Catch and Gutter. For the ball catching task, the agent was trained with the goal of catching a ball that is thrown toward the agent (robot). Regarding the Gutter task, the agent was trained to place a ball in a desired position of the gutter. In both tasks, agents need to be provided with information about the environment to perform the tasks.

Training agents in complex tasks usually requires a considerable amount of trial and errors. A solution to decrease the training time required is to use Transfer Learning. This method is a promising technique in the field of ML that aims to reduce the time and costs associated with training robots. Although this approach works well in simulations, it faces challenges when transferred to real and to others robots. In addition, in the implementation of dynamic tasks, there are often issues as the trained agent requires frequent information about the environment. Therefore, to address the issue of delayed information provided to the agent, solutions with low latency need to be implemented.

Usually one of the sensors used to provide information to agents are cameras that provide frames. However, for dynamic tasks this technology has its limitations. Such limitations are linked to the information rate it can provide (usually 30fps). Furthermore, it should be taken into consideration the time required to process the information provided by the camera, making it an obsolete option for the application. On the other hand a recent technology that has gained a significant application in the field of Computer Vision are the event based cameras. Such technology has a low latency in the range of us, making it an excellent option to address the issue of delayed information to the agent. In this context, this internship involves the application of an event based camera to detect the position of the ball for the execution of dynamic tasks. The results and insights obtained during the internship will be used for the development of this part of the thesis.

To validate the transfer methodology developed during the PhD, a total of 3 tasks have already been performed in simulation, namely: Peg Insertion 1.0.1a, Pick and Place 1.0.1b, and Ball Catching 1.0.1c. All three tasks have already been developed and are functional in simulation by Mr. Samuel Beaussant. The Peg Insertion and Pick and Place tasks have already been performed with the real robots. Thus, the contribution of this internship is to perform in real robots the tasks Ball Catching and Gutter using a event based camera to provide information to the agent about the environment.



(a) The Peg Insertion task with the 7 DoF Panda robot.

(b) The Pick and Place task with the 6 DoF UR10 robot.

(c) The Ball Catching task with the 6 DoF Panda robot.

Figure 1.0.1: The three tasks considered to validate the transfer approach in Unity.

The overall objective of this internship is to perform ball detection using an event based camera for the ball catching and Gutter tasks, which will be carried out with a robot. The specific objectives are as follows:

- Investigate and understand the principles of Computer Vision applied to event based cameras,
- Implement a ball detection system using an event based camera,
- Apply RL algorithms and transfer learning methods to train robotic agents for specific dynamics tasks.

The report begins with a description of the company/organization, highlighting its research areas, activities, collaborations, and organizational structure. The next section focuses on the theoretical framework, exploring concepts such as Machine Learning, Transfer Learning, and Computer Vision. The report then delves into the development of the internship, discussing specific tasks performed, such as Ball Detection, the Gutter task, and the Ball Catching task. This is followed by a section on cultural differences, which discusses aspects such as climate, cuisine, and cultural and professional aspects in the workplace. Finally, the report concludes by summarizing the main findings, contributions, and potential future work.

CHAPTER 2

Company/Organization Description

In this chapter, I will present the institution where I did my internship. First, I will provide a brief introduction about the Pascal Institute and its partnerships, then I will show the organizational structure. Finally, I will introduce the team in which I was incorporated during the internship.

2.1 Company Overview

The Institut Pascal is an interdisciplinary research and training unit located at Université Clermont Auvergne (UCA) and under the supervision of CNRS (Centre National de la Recherche Scientifique). Comprising approximately 400 individuals, the institute is considered a leading research unit in the region [2].

Figure 2.1.1: Insitut Pascal[1].

2.2 Research Areas and Activities

The Institut Pascal came together after the merge of seven laboratories at different times (2012, 2017, and 2021), with the aim of structuring research in the fields of Engineering Sciences and Systems on the Clermont campus. These areas include Process Engineering, Mechanics, Robotics, Physics of Information Sciences, and Health.

[1] www.institutpascal.uca.fr

Within these areas, the institute develops knowledge and technologies with applications in three main domains: factories (including ecosystems), transportation, and hospital of the future. These areas reflect the Institut Pascal's commitment to seeking innovative solutions to contemporary challenges.

2.3 Collaborations and Partnerships

The Institut Pascal is a member of FACTOLAB, a joint laboratory established in collaboration with the company MICHELIN. Additionally, the institute leads the IMobS3 laboratory of excellence and is a member of the CNRS EquipEx ROBOTEX network, as well as the LabEx GaNeX (PIA1) and PRIMES networks. The unit also actively participates in partnerships with the competitiveness clusters CIMES, AXELERA, MINALOGIC, POLYMERIS, and XYLOFUTUR, through a partnership with Université Clermont Auvergne. The Institut Pascal is also a member of the renowned Carnot Institute MECD.

2.4 Organizational Structure

As shown in Figure 2.4.1, the Institu Pascal is organized into five research groups.

- Process Engineering, Energy and Biosystem (GePEB). This axis caters to a wide range of application domains, including the generation of energy carriers, the production of biomaterials, the enhancement of food processing, and various other areas,
- Mechanics, Mechanical Engineering, Civil Engineering, Industrial Engineering (M3M). It is a multidisciplinary research area that strives to address the scientific challenges presented by industrial needs in the domains of mechanics and civil engineering,
- Image, Perception Systems, Robotics (ISPR). This axis works in the field of Artificial Perception and Vision for the Control of Robotic Systems. Its objective is therefore the development of theoretical, methodological and architectural concepts for the perception and control of systems,
- Image Guided Therapies (TGI) brings together 4 research themes: Cardio-Vascular Interventional Therapy and Imaging, Endoscopy and Computer Vision, Image-Guided Clinical Neuroscience and finally Perinatal,
- And finally, Photonics, Waves, Nanomaterials (PHOTON) is particularly interested in nanophotonics, nanostructures, microsystems and nanomaterials and electromagnetic compatibility.

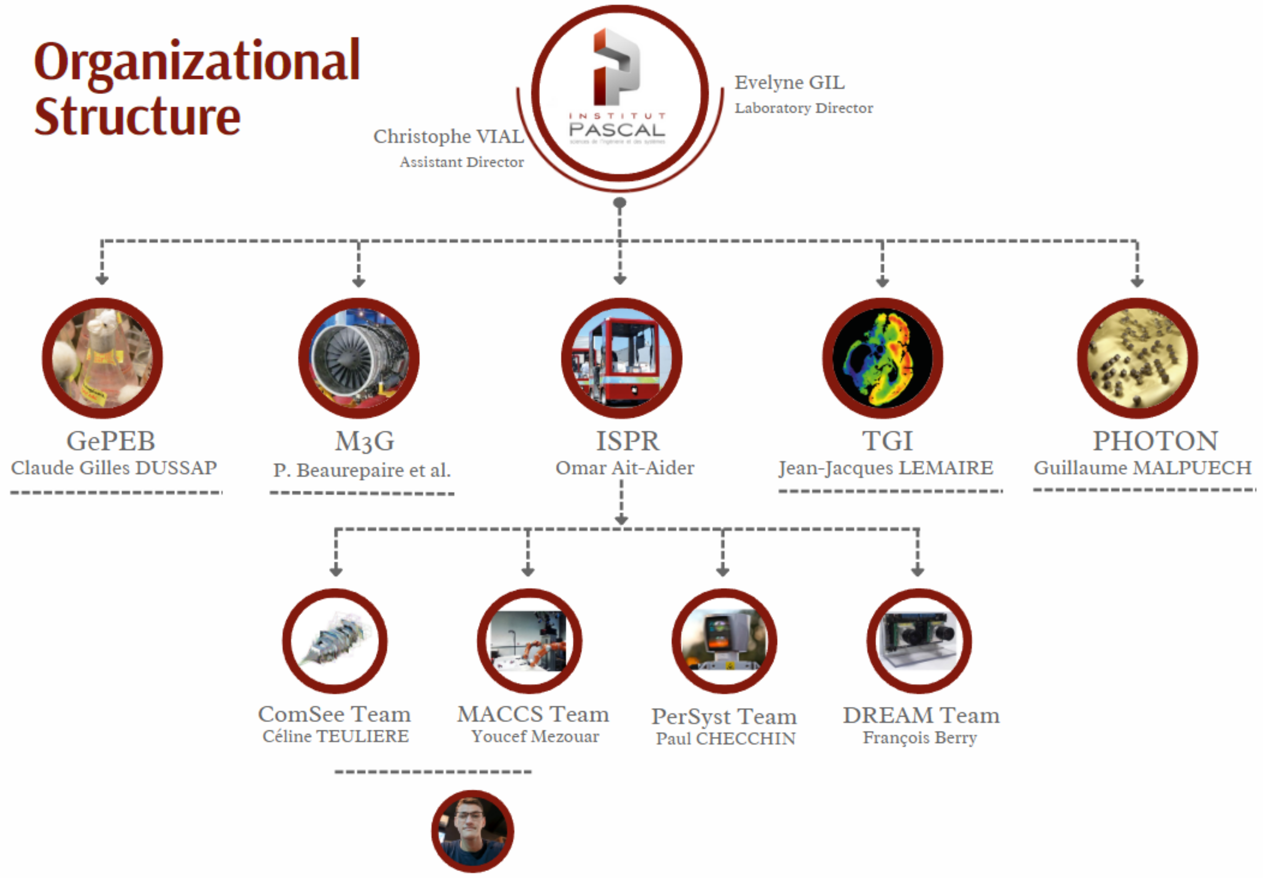


Figure 2.4.1: Organizational structure Insitut Pascal.

2.5 Team and Research Laboratory

In my internship, I was part of the ISPR axis under the supervision of Mr. Sebastien Lengagne and Mr. Samuel Beaussant. This axis is divided into smaller research units:

- Artificial Vision (ComSee),
- Modeling, identification and command (MACCS),
- Multi-sensory perception systems (PerSyst),
- Hardware and software architecture for perception (DREAM).

I was working with the MACCS team (Modeling, Autonomy and Control in Complex Systems). Furthermore, I had a significant interaction with the ComSee team (artificial Vision) in which Ms. Céline Teulière and Mr. Antony N'Dri served as my tutors in the Computer Vision part.

CHAPTER 3

Theoretical Framework

The goal of this chapter is to provide an overview of the work done by Mr. Samuel Beaussant with respect to transfer learning and explain the operation of and differences between an event based camera and a frame based camera. First we introduce basic concepts of ML and its sub-domains. Then we explain the transfer methodology developed during the PhD. Finally we present the event based camera and its characteristics.

3.1 Machine Learning

ML is a subfield of artificial intelligence that focuses on enabling computers to learn from datasets and learning algorithms. Learning, in this context, refers to a computer's ability to improve its performance on one or more tasks autonomously, without explicit programming. To achieve this, the algorithm constructs a model based on the data available in a dataset. The model is then used to make predictions when new data is presented to it. The primary sub-domains of ML there are Supervised Learning, Unsupervised Learning and RL. Usually for solving ML problems we use one of these three methods depending on the types of data available.

3.1.1 Supervised Learning

According to [3] supervised learning is the field of ML in which a model is trained on a labeled dataset, where each training example has an associated correct label or response. The goal of supervised learning is for the model to learn how to make accurate predictions for new examples not previously seen, based on the information provided by the training data and their corresponding labels. An example of supervised learning is training a model for image classification. In this case, the dataset consists of labeled images, where each image is assigned a class such as "cat" or "dog" as shown in Figure 3.1.1. The supervised learning algorithm is trained using these images and their corresponding labels. After training, the model can classify new images, determining whether a given image is of a "cat" or a "dog" based on the patterns and features learned during training.

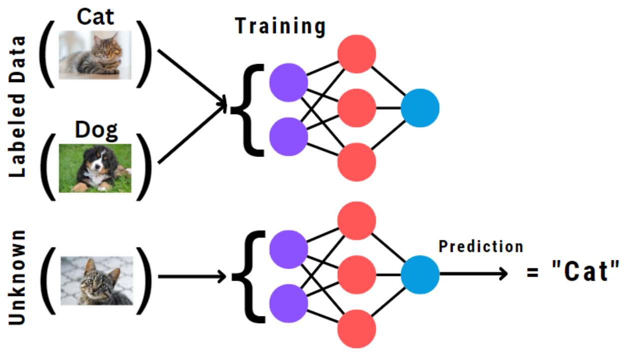


Figure 3.1.1: An example of image classification in Supervised Learning.

3.1.2 Unsupervised Learning

In contrast to supervised learning, where the data is labeled, unsupervised learning does not have information about which class each item in the dataset belongs to. Therefore, in unsupervised a model is trained on a set of unlabeled data, i.e., data without labels or known correct answers. According to [3] the goal of unsupervised learning is to discover structures, patterns, or relationships hidden in the data, without the guidance of labels. This usually involves grouping the data into clusters with similar characteristics or reducing the dimensionality of the data to facilitate analysis. Differently from the example of Supervised Learning, in Unsupervised Learning, as shown in Figure 3.1.2, the model is trained with an unlabeled dataset. Once these data are passed to the algorithm, it tries to find patterns to separate them into groups based on their characteristics. In our example, the algorithm separates the data into three groups: "cats", "dogs", and "lions".

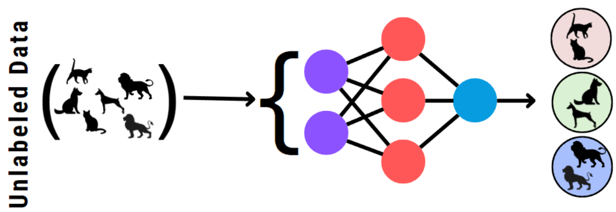


Figure 3.1.2: An example of clustering with unsupervised Learning.

3.1.3 Reinforcement Learning

According to [4], RL involves learning how to make sequential decisions in an environment to maximize the cumulative reward. As shown in 3.1.3 the agent interacts with the environment, observing states, taking actions and receiving rewards, which can be positive or negative. The goal of RL is to learn an action policy that allows the agent to make sequential decisions that lead to the most rewarding actions in different states of the environment.

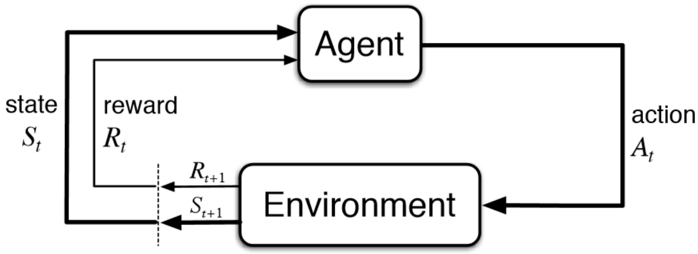


Figure 3.1.3: Reinforcement learning feedback loop[1].

A notable example of RL is AlphaGo, a computer program developed by DeepMind [5]. AlphaGo was trained to play the board game Go, which is considered extremely complex due to the large number of possible moves. This model achieved a significant milestone by defeating the world champion of Go in a series of games in 2016, demonstrating the capability of RL to master a complex and challenging game. For more information, a documentary is available here.

3.2 Transfer Learning

Transfer Learning aims to enable an artificial intelligence model to apply acquired knowledge from a specific task or robot to enhance its performance in related new tasks. This field aims to achieve the ability to transfer skills from one context to another, thereby reducing the need for intensive training in each new task.

To promote the Transfer Learning between agents, it is important to explore advanced techniques. In the PhD thesis of Mr. Samuel Beaussant entitled "Transfer Learning in Robotics with State Abstraction" he exploits Transfer Learning in the context of RL for the specific purpose of transferring policies from one agent to another, even in the presence of morphological discrepancies or different state action spaces.

This approach involves creating a shared latent space where representations of related tasks are mapped. In this context "Latent Space" refers to an abstract multidimensional space containing values of characteristics that we cannot interpret directly, but that encode a representation of the dosed state.

The method implemented by Mr. Samuel Beaussant in his PhD thesis can be divided into three modules (See Figure 3.2.1), namely: "Input Base", "Task module" and "Output Base". The information provided by the environment are divided into two types, Robot-specific information and task-specific information. Thus, task-specific information fed the Universal Notice Network (UNN). As proposed in [6], the objective of the UNN module is to provide a dedicated task module that can be set in the middle of a control pipeline, allowing any compatible robot to solve the task. The robot-specific information feeds the "Input Base" module, which is responsible for mapping the robot space to the shared resource space where UNN operates. Finally, the "Output Base" module remaps the actions generated by UNN to the specific robot space.

[1]towardsdatascience.com

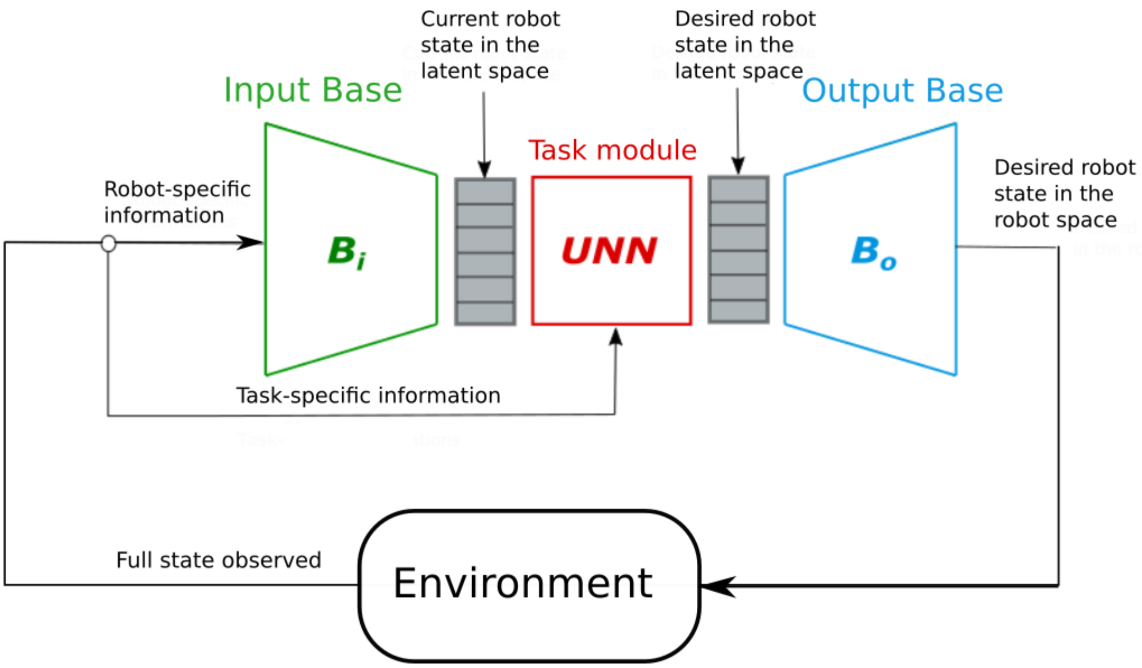


Figure 3.2.1: Latent space UNN pipeline.

3.3 Computer Vision

In this section, we will present the relevant information regarding vision and image processing for carrying out dynamic tasks. The camera used was a DAVIS 346 shown in Figure 3.3.1, which is equipped with a DVS sensor of 346×260 pixels. For more information you can view the Appendix A.

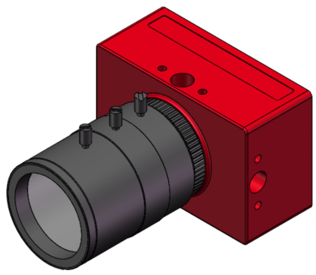


Figure 3.3.1: Event based camera used in the internship[2].

An event based camera is a type of camera that captures individual events instead of complete frames, using an asynchronous approach based on changes in pixel brightness. These cameras are highly sensitive to rapid changes and can capture events on a us timescale. As shown in [7], event cameras have several applications involving fast motion, object tracking, and Computer Vision.

To better understand how event based cameras work, we will show how they work by comparing them to frame-based cameras. As an example, we will consider a disk that has a black circle near the edge. We will look at three cases: when the disk is moving at a low speed, when the disk is not moving, and finally when the disk is moving at a high speed.

[2]grabcad.com

We can see in Figure 3.3.2 that when the disk is in motion, the black circle generates a luminosity variation in the camera sensor. In this way, events are generated, containing information such as the position relative to the sensor (x, y), the time at which the event occurred (*timestamp*), and the polarity that indicates whether there was a brightness increase (+1) or decrease (−1). We can notice that the rest of the circle that has a homogeneous color does not generate events because it is not causing a luminosity. For a frame-based camera we have as results frames with defined intervals thus causing a significant loss of information of the position of the black circle during this time interval.

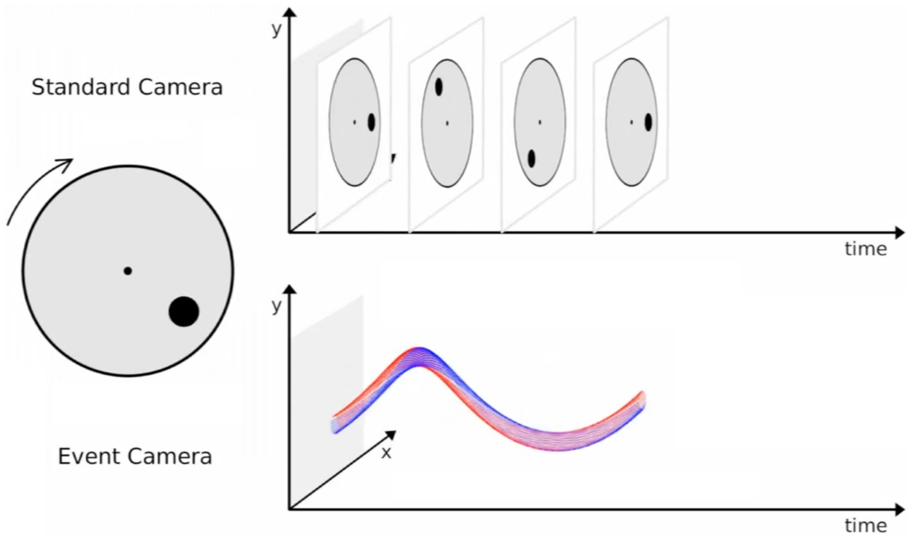


Figure 3.3.2: Visualization of the output from an event camera sensor and a frame-based camera when facing a rotating disc with a black circle. Adapted from here.

In contrast, as shown in Figure 3.3.3 when the disk is not in motion, no information is provided by the event camera. However, for a standard camera, we still have the all the frames recorded. This characteristic prevents the event camera from producing redundant data.

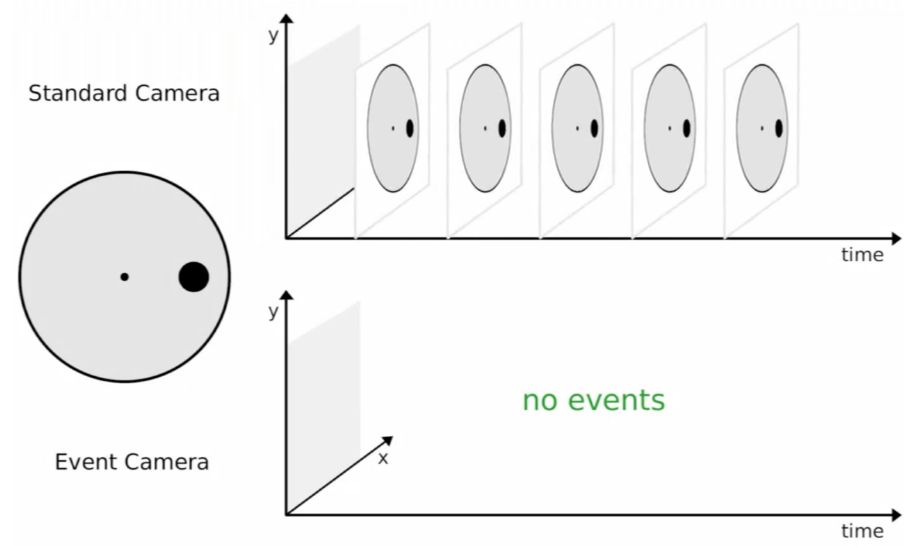


Figure 3.3.3: Visualization of the output from an event camera sensor and a frame-based camera when the disc is without movement. Adapted from here.

Finally, when the disk is rotating at a high speed (see Figure 3.3.4), the event camera can provide a good resolution of the circle's position, while the standard camera produces a blurred image. The blurred effect on frames captured by frame-based cameras occurs due to the exposure time required to capture the frame. When a moving object is photographed, the camera needs to keep the shutter open for a certain period to capture the necessary light. During this time, if the object is moving, its position will change, resulting in a shift of the recorded image relative to the object.

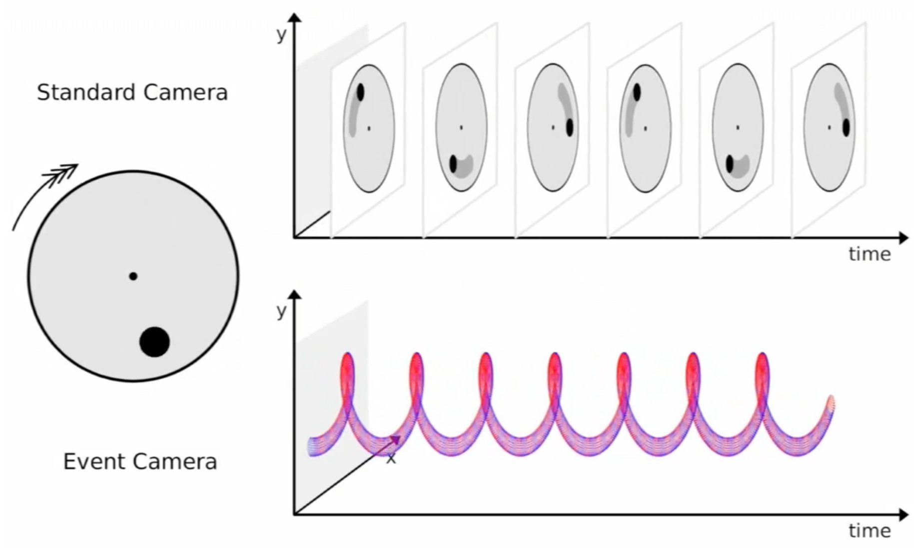


Figure 3.3.4: Visualization of the output from an event camera sensor and a frame-based camera when the disc is moving fast. Adapted from here.

CHAPTER 4

Internship Development

This chapter presents the work carried out during the internship. The activities can be divided into three sections. The first section encompasses all the development related to Computer Vision and used during the tasks considered. The second section (Gutter Task) involves the modeling in simulation and implementation on real robots of a task previously performed in [8]. The last section refers to the implementation of results obtained in simulation by Mr. Samuel Beaussant on the real robots for the Ball Catching task.

4.1 Ball Detection

In this section we will present the work carried out related to detecting the ball with an event based camera. First we present a brief description of the problem. Then we show the proposed solution using a clustering algorithm to detect the ball. Finally, we show the results obtained to detect the ball in the tasks Gutter Task and Ball Catching Task. This part of the internship is crucial, since it is the main contribution to Mr. Samuel Beaussant PhD thesis.

4.1.1 Description

For ball detection, as mentioned before, the goal is to use an event based camera. In the Figure 4.1.1 we can see an example of an image generated by the accumulation of events during a short period of time. Once the data is received the main problem becomes to know which group of events is the ball. With a visual analysis we can see that the ball when thrown has a high density of events grouped when compared to the other objects. In this way, we can use a density-based grouping method to group the events that belong to the ball and then filter it.

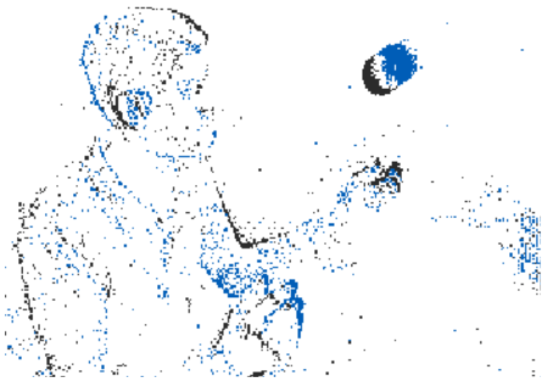


Figure 4.1.1: Accumulated events over a short period of time.

In the Gutter task the detection of the ball needs to be done with a precision less than 2 *cm*. For the Ball Catch task the robot's effector will be a basket with a radius of 12 *cm* and the ball thrown towards the robot has a radius of 4 *cm*, so we will have a margin of error of 8 *cm*. in both tasks the robots need to be fed with the information every 20 *ms*. In the Figure there is the table of specifics for each task.

Ball Detection in Gutter Task		
Requirements	**Criterion**	**Description**
Detect position of the ball	accuracy	Accuracy detection < 2 cm
Time detection of the ball	Time	Time detection < 20 ms
Ball Detection in Ball Catch Task		
Requirements	**Criterion**	**Description**
Detect position of the ball	accuracy	Accuracy detection < 8 cm
Time detection of the ball	Time	Time detection < 20 ms

Figure 4.1.2: Terms of Reference.

4.1.2 Proposed solution

As the data provided by the event based camera has coordinates (x,y), we can use these coordinates to create different groups of objects that are being detected by the camera in order to have the information of which object is the ball. So, to separate the data into different groups of objects the proposed solution involves the application of the density based clustering algorithm DBSCAN demonstrated in [9]. DBSCAN is a widely used algorithm in clustering tasks because it can group data based on spatial density. In this context, the events captured by the camera are considered as spatial points of two dimensions (x,y), these coordinates are used as data in DBSCAN.

This approach allows for fast and reasonably accurate detections, even with the asynchronous nature of the events captured by the camera, making it suitable for the tasks. Object detection using event based cameras and DBSCAN has been previously implemented in [10].

In addition to coordinates data, DBSCAN clustering algorithm requires others two parameters:

- **minPts:** The minimum number of data points required for a cluster to be considered a dense region,
- **eps:** The distance measure used to define the neighborhood of a data point. If the distance between two data points is less than or equal to eps, they are considered neighbors.

Therefore, based on these two parameters, the DBSCAN algorithm categorizes points into three types (See Figure 4.1.3):

- **Core Points:** A data point is classified as a core point (red points in Figure 4.1.3) if there are at least **minPts** number of data points within the radius of **eps**,
- **Border Points:** A data point is considered a border point (gray points in Figure 4.1.3) if it is in the neighborhood of a core point but has fewer than **minPts** data points within the radius of **eps**,
- **Noise Points:** A data point is considered a noise point (yellow points in Figure 4.1.3) if it is neither a core point nor in the neighborhood of a core point.

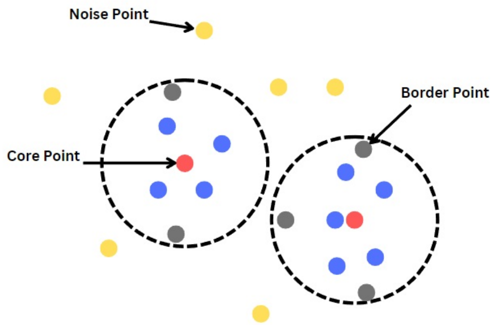


Figure 4.1.3: Core, Border and Noise points of the DBSCAN.

Once the data is grouped into their respective clusters with DBSCAN, we need to filter the cluster that is the ball. To filter the ball we can make an analysis of the events produced by a ball when placed in motion in front of an event based camera. In our case for both tasks the object to be detected is a ball. Therefore, we have a circular object. In Figure 4.1.4a, we can see a diagram of a circle moving horizontally. In this scenario, the background is darker than the ball. In this diagram, we have positive events generated on the right side of the circle and negative events generated on the left side.

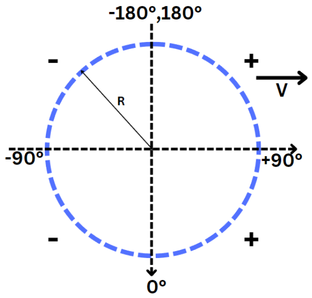
As shown in Figure 4.1.4b, the events generated due to the ball moving at considerable speed do not form a perfect circle. Instead, more events are generated at the edges of the direction of movement. The higher the speed, the closer the generated events resemble an ellipse.

We can see in Figure 4.1.4c, the distribution of events at different distances from the center. The center of the circle (red cross) was calculated using the average of the events, as demonstrated in equation 4.1.1. Once the center was found, the distance was calculated in relation to this central point using equation 4.1.2. In this distribution (see Figure 4.1.4c), it can be observed that the mean is close to the radius of the circle. This distribution also exhibits low variance.

$$(\bar{x}, \bar{y}) = \frac{\sum_{i=1}^{n} (x_i, y_i)}{n} \tag{4.1.1}$$

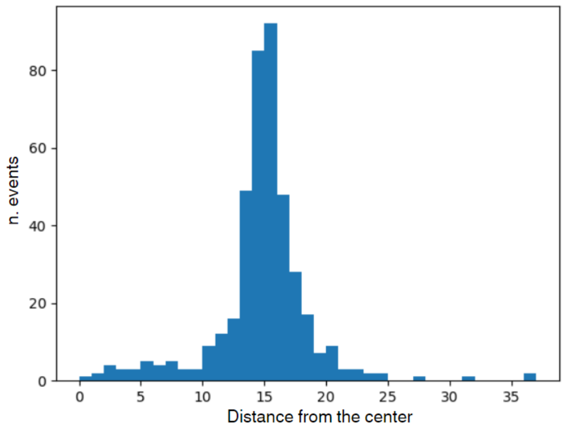
$$d_{event,center} = \sqrt{(x_i - \bar{x})^2 + (y_i - \bar{y})^2} \tag{4.1.2}$$

When calculating the angles of events (See Figure 4.1.4d), we can observe that they have a bimodal distribution. In our case to solve the problem of filtering the desired cluster, we used the variance of distances and angles as parameters.

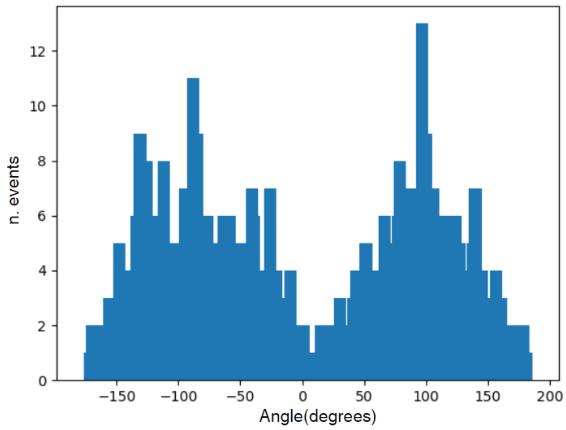


(a) A moving circle.

(b) Information provided by the circle movement.



(c) Distribution of all events over different distances from the origin.



(d) Distribution of all events over different angles.

Figure 4.1.4: Analysis of a circular object moving in front of an event based camera. Inspired by [11].

4.1.3 Results

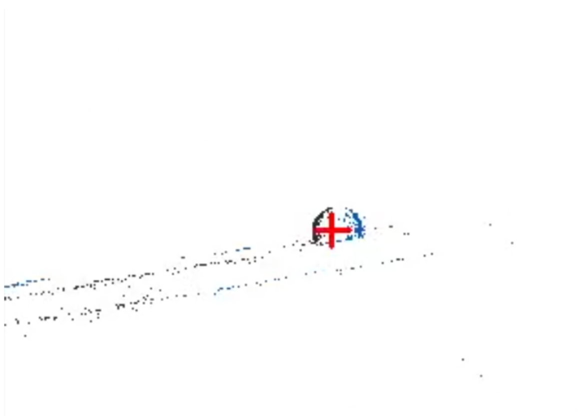
In the ball detection, two different codes were implemented. The first algorithm was developed specifically for the interaction with the gutter, while the second one was designed for the Ball Catching task.

4.1.3.1 Detection Ball in Gutter Task

In the Gutter task, at the beginning, both the ball and the robot are in a resting state, which means there are no events captured by the event based camera. However, the used camera has the capability of providing both events and grayscale frames. Therefore, to solve this boot issue, the images provided by the camera were used.

Once we have the grayscale image, the "Circle Hough Transform" algorithm was applied to the frame, aiming to detect circular geometric shapes (ball). This approach allowed for the identification of the ball's presence even when there are no recorded motion events, facilitating the detection process at the beginning of the task execution. The initialization image is shown in Figure 4.1.5a and the figure 4.1.5b shows the position of the ball (red cross) found with the events. Videos are available for viewing here.

(a) Initial detection with frames.



(b) Detection using events provided by the camera.

Figure 4.1.5: Ball detection in the Gutter task.

4.1.3.2 Detection Ball in Ball Catching Task

For the Ball Catching task, unlike the Gutter task, only events were used. In Figure 4.1.6, we can see the detection of the ball being thrown by a person. In this case, we can observe the complexity of the detection. The environment includes a person and various events generated by noise, however, even with the complexity, the algorithm achieved good performance by using variance to filter the desired cluster. It should be noted that, for the implementation of the Ball Catching task, we will not have the same complexity in relation to the generated events, because the person throwing the ball to the robot will not be present in the camera's detection field. Videos are available for viewing here.

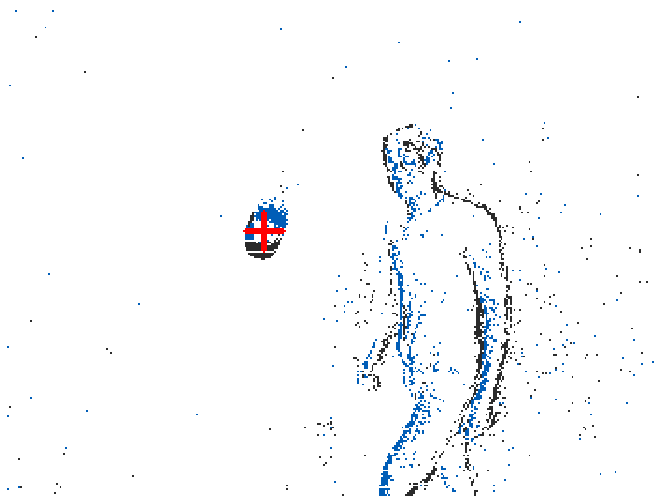


Figure 4.1.6: Ball detection in the Ball Catching task.

The trained agents require environment information every 20 ms for proper task performance. In the specific case of the Gutter task, the environment information is provided at intervals that can vary between 6 ms and 11 ms. This update frequency is within the established time limit, ensuring that the agents receive the necessary information to perform actions and make decisions appropriately.

4.2 Gutter Task

This task was previously performed by Mr. Samuel Beaussant in [8]. On that occasion, he carried out the task with two robots, both trained with delay aware and delay unaware in the training environment. One of the objectives of the publication was to demonstrate an improvement in the robot's performance when trained with awareness of the delay. This is due to the fact that the information provided to the agent (robot) is obtained from a frame-based camera, causing a delay between the actual state of the environment and the robot's perception. The objective of redoing this task is to demonstrate whether using an event based camera can result in better robot performance in performing the task compared to performance with frame-based cameras, as we won't have the issue of delayed information provided to the agent.

4.2.1 Description

The environment shown in Figure 4.2.1 consists of a gutter with one end (right side) fixed and a robot responsible for moving the gutter at the other end. In this task a robot needs to keep a ball (blue) at a desired position (green) on a gutter. The main differences compared to the implementation in [8] are related to the agents and the timing at which information is provided to them. In our case, we use the Panda and UR10 robots, which have different sizes and numbers of joints. Additionally, the camera used to provide information to the agent will be an event based camera.

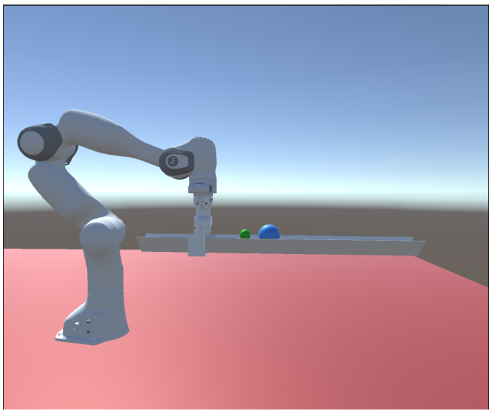


Figure 4.2.1: Gutter task training environment with the Panda robot.

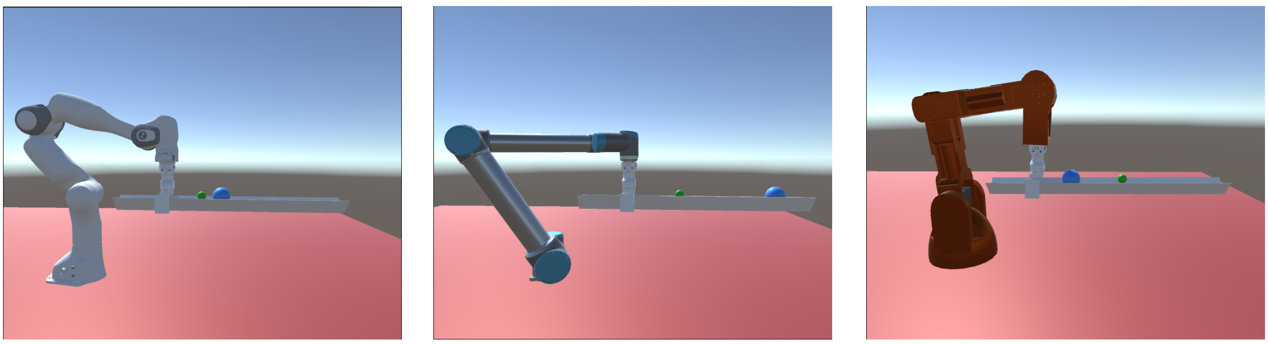
4.2.2 Training

Training was realized with three different robots as shown in Figure 4.2.2. The execution of the task in the real world will be performed only with the Panda robot (see Figure 4.2.2a) and the UR10 robot (see Figure 4.2.2b), as the Braccio robot (see Figure 4.2.2c) is a small robot that cannot perform the task. Through the previously described operational dynamics, this task can be formalized with the following Markov Decision Process (MDP):

- **State:** $s_t \in \mathbb{R}^{6+2n}$: the ball position and velocity on the gutter, the desired ball position, the effector position (x, y, z), the velocity and position of robot joints (n being the number of considered joints).
- **Action:** $a_t \in \mathbb{R}^n$ is the velocity command of the joints (n being the number of considered joints).
- **Reward:**

$$r_t = \begin{cases} r & if \quad d_{des,b} < \delta \\ 1/\alpha d_{des,b} & else \end{cases} \tag{4.2.1}$$

where r represents a small positive reward, δ is the reward area of a small reward, $d_{des,b}$ denotes the distance between the ball and the desired ball position and α a weighting constant.



(a) 7DoF Panda robot. (b) 6DoF UR10 robot. (c) 5DoF Braccio robot.

Figure 4.2.2: Robots used in the training of the Gutter task.

For training the agents, Proximal Policy Optimization (PPO) [12] was used. PPO is a widely used optimization algorithm in RL. It belongs to the class of gradient ascent algorithms, which aim to find the optimal policy for a reinforcement learning agent. The goal of PPO is to improve an agent's policy iteratively, seeking to maximize the accumulated reward over time. The main feature of PPO is its policy update approach, which aims to balance environment exploration and learning stability.

4.2.3 Results

The task involving the Gutter consisted of training the agents, which is currently ongoing with the use in the real robots. At the current stage of the internship, the agents are being tested and evaluated in the real robots, aiming to improve their performance ability in the task. As shown in Figure 4.2.3 all agents converged to a good reward value. Of the three agents who got a better average reward was the Panda robot. It should be noted that the three robots have different morphologies (number of Joints). Videos showing the robots performing the task in simulation can be found here.

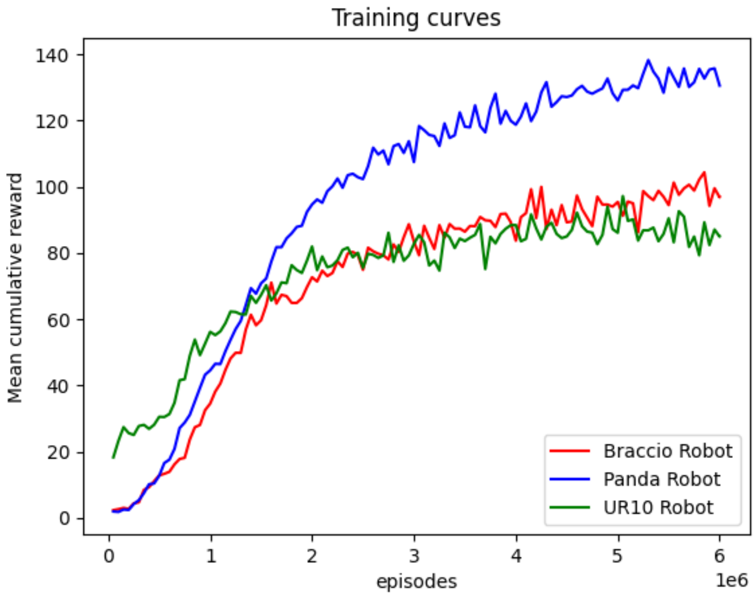


Figure 4.2.3: Training curves. All the agents were trained for 6000000 steps.

To perform manipulations with real robots, the environment in which the task is being executed can be visualized in Figure 4.2.4. Based on the results obtained so far, it has been found that the model needs to undergo further training. This necessity arises from the differences in dynamics between the ball and the gutter, which diverge from the dynamics observed in the simulation. This discovery highlights the importance of adapting and adjusting the control and learning models to deal with the specificities of the real world, overcoming the specific challenges encountered during the execution of the task in physical environments.

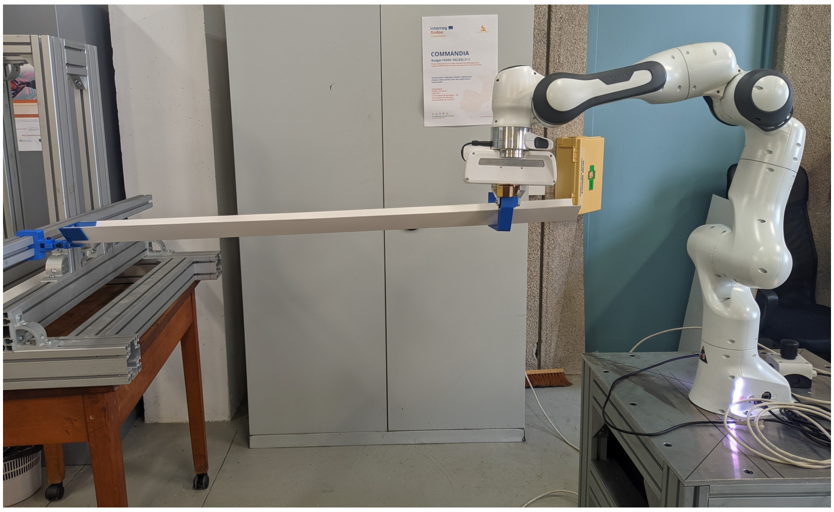


Figure 4.2.4: Environment for performing the Gutter task with the Panda robot.

4.3 Ball Catching Task

For the Ball catching task shown in Figure 4.3.1, the robot training had already been conducted by Mr. Samuel Beaussant is missing only the implementation of the task in the real robots. Therefore, all the information mentioned here in this section has as reference your thesis.

4.3.1 Description

In this task, a basket is placed as the robot's effector, and the agent's goal is to catch a thrown ball before touching the ground. The ball is thrown from the same location, but the trajectory varies by randomly selecting the velocity of the ball along the three spatial axes (x,y,z). This task serves to demonstrate the ability of UNN to handle dynamic tasks.

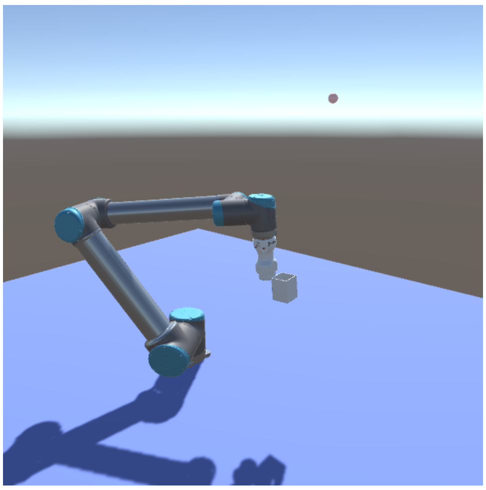


Figure 4.3.1: Ball Catching task training environment with the UR10 robot.

4.3.1.1 Training

The environment consists of a robot with a basket. The objective is to prevent the ball thrown towards the robot from touching the ground. In order for the agent to successfully complete the task, the position of the ball is provided.

- **State:** $s_t \in \mathbb{R}^{13+2n}$: the ball position (x,y,z) and velocity (x,y,z), the basket position (x,y,z) and basket velocity (x,y,z), a flag catch to indicate to the agent if the ball is inside the basket, the velocity and position of robot joints (n being the number of considered joints).
- **Action:** $a_t \in \mathbb{R}^n$ is the velocity command of the joints (n being the number of considered joints).
- **Reward:**

$$r_t = \begin{cases} -c.d_{b,t} - \beta|\theta_e| & if \quad catch = True \\ e - \beta|\theta_e| & else \end{cases} \tag{4.3.1}$$

  where $d_{b,t}$ is the distance between the basket and the target receiving position. The constant c is a small scaling constant and e is a small positive constant reward. In your settings, $c = 0.15$ and $e = 0.15$. The term $|\theta_e|$ is the angle between the effector pose and the vertical plane which guide the agents towards suitable body configuration.

4.3.2 Results

At the time of writing, we only have simulation results for the Ball Catching task. Videos showing the performance of the robots in simulation for this task can be found here.

CHAPTER 5

Cultural Differences

During my internship, I had the opportunity to experience and learn about the cultural differences between Brazil and France. These differences were noticed in various aspects of daily life, including cuisine, traditions, and especially the climate.

5.1 Climate

Brazil has a predominantly tropical climate, with warmer temperatures and high humidity for a significant part of the year. In my city (São Luís, MA), the average maximum annual temperature is around 34 degrees Celsius with a heat index of 38≈40. The average minimum temperature is approximately 24 degrees Celsius. As a result, we do not have well-defined four seasons, with more noticeable contrasting periods of winter and summer.

In France, the four seasons are well defined, allowing for a better perception of the changing seasons. In the city where I did my internship (Clermont-Ferrand), I had the opportunity to witness temperatures that reminded me of Brazil. However, I also experienced low temperatures, which allowed me to have one of the best experiences, which was getting to know snow.

5.2 Cuisine

French cuisine is internationally renowned for its sophistication and appreciation of fresh, high-quality ingredients. The French have a strong appreciation for the art of cooking and well-prepared meals, where the presentation of dishes is as important as the taste.

On the other hand, Brazilian cuisine is known for its diversity and influences from different cultures, such as Indigenous, African, and European. Brazilian food is rich in intense flavors and varied spices.

Some popular dishes from Brazilian cuisine include:

- Feijoada: A traditional Brazilian dish made with black beans, pork, sausage, and other ingredients, served with rice, collard greens, and orange slices.
- Acarajé: A typical delicacy from Bahia, made with fried black-eyed pea dough and filled with vatapá, dried shrimp, and other delicacies.

- Pão de Queijo: A Minas Gerais specialty made with a cheese and cassava flour dough, baked until crispy on the outside and soft on the inside.

During my time in France, I had the opportunity to experience new flavors that differ in several aspects from Brazilian cuisine. These culinary differences reflect the traditions, history, and unique characteristics of each country.

5.3 Cultural and Professional Aspects in the Workplace

Brazil and France present significant differences regarding the work environment. In Brazil, the culture is generally more informal, and this informality has its advantages up to a certain extent. However, Brazilians have the habit of arriving late, which, in my opinion, is not a good practice. In the meetings I attended during my internship, this was one of the aspects that I appreciated: the fact that people arrived on time and were effective throughout the meeting.

These differences reflect the cultures and values of each country. While Brazil values informality, France prioritizes formality and structure in the workplace. It is important to understand and respect these differences.

CHAPTER 6

Conclusion

This report presented the applications of RL and Transfer Learning techniques for performing dynamic tasks with robots. Additionally, it focused on the area of Computer Vision, where a system was implemented using an event based camera to detect the position of a ball. The main objective of this internship was to develop an efficient method for ball position detection and contribute to the research on transfer learning in the PhD thesis of Mr. Samuel Beaussant.

Currently, the results obtained regarding ball detection have been promising. Regarding the Gutter task, the specifications related to detection time and accuracy have been met. As for the Ball Catching task, approximately 70% of the process has been implemented up to the present moment. The detection time has also been respected, with the need for further improvement in the accuracy of one of the ball's position coordinates.

Regarding the implementation of the Gutter task on real robots, the models are undergoing a new training phase. However, based on the results already obtained, it has become evident the complexity of implementing these tasks on real robots. This additional training step aims to improve the performance and precision of the robots, adapting them to the specificities of the real environment.

The internship is still in progress, and as future work, there is the remaining task of completing the implementation of the Gutter task and Ball Catching task with real robots.

During my internship in Institut Pascal, several competences essential in the field of research and development in robotics were acquired. One of the main skills acquired was the familiarity and application of fundamental concepts in ML, transfer learning, and reinforcement learning.

Furthermore, the internship provided an opportunity to develop practical skills in programming, implementing ML algorithms and Computer Vision.

Finally, the internship also helped to develop communication and collaboration skills. Participation in team meetings, presentations, and technical discussions allowed for the exchange of knowledge and experiences with other team members.

Bibliography

[1] OpenAI: Marcin Andrychowicz, Bowen Baker, Maciek Chociej, Rafal Jozefowicz, Bob McGrew, Jakub Pachocki, Arthur Petron, Matthias Plappert, Glenn Powell, Alex Ray, et al. "Learning dexterous in-hand manipulation". In: *The International Journal of Robotics Research* 39.1 (2020), pp. 3–20.

[2] *Instut Pascal*. `http://www.institutpascal.uca.fr/index.php/en/`. Accessed on 24 June 2023. 2023.

[3] C. M. Bishop. *Pattern Recognition and Machine Learning*. New York, NY: Springer, 2006.

[4] Richard S. Sutton and Andrew Barto. *Reinforcement Learning: An Introduction*. New York, NY: MIT Press, 1998.

[5] David Silver, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, et al. "Mastering the game of Go with deep neural networks and tree search". In: *nature* 529.7587 (2016), pp. 484–489.

[6] Mehdi Mounsif, Sebastien Lengagne, Benoit Thuilot, and Lounis Adouane. "Universal notice network: Transferable knowledge among agents". In: *2019 6th International Conference on Control, Decision and Information Technologies (CoDIT)* (2019), pp. 563–568.

[7] Guillermo Gallego, Tobi Delbrück, Garrick Orchard, Chiara Bartolozzi, Brian Taba, Andrea Censi, Stefan Leutenegger, Andrew J. Davison, Jörg Conradt, Kostas Daniilidis, and Davide Scaramuzza. "Event-Based Vision: A Survey". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44.1 (2022), pp. 154–180.

[8] Samuel Beaussant, Sébastien Lengagne, Benoit Thuilot, and Olivier Stasse. "Delay Aware Universal Notice Network: Real world multi-robot transfer learning". In: *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (2021), pp. 1251–1258.

[9] Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu, et al. "A density-based algorithm for discovering clusters in large spatial databases with noise." In: *kdd* 96.34 (1996), pp. 226–231.

[10] Benedek Forrai, Takahiro Miki, Daniel Gehrig, Marco Hutter, and Davide Scaramuzza. "Event-based Agile Object Catching with a Quadrupedal Robot". In: *IEEE International Conference on Robotics and Automation (ICRA)* (2023).

[11] Sajjad Seifozzakerini, Wei-Yun Yau, Kezhi Mao, and Hossein Nejati. "Hough transform implementation for event-based systems: Concepts and challenges". In: *Frontiers in computational neuroscience* 12 (2018), p. 103.

[12] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. "Proximal policy optimization algorithms". In: *arXiv preprint arXiv:1707.06347* (2017).

Appendices

APPENDIX A

DataSheet event based camera DAVIS346[1]

[1] inivation.com

DAVIS 346

The DAVIS 346 is a 346 x 260 pixels DVS event camera with included active pixel frame sensor.

Specifications

DVS Resolution	346 x 260 pixels
Frame Resolution	346 x 260 pixels, Grayscale, simultaneous output with DVS
DVS Dynamic range	120 dB
APS Dynamic range	56.7 dB
Min. latency	~ 20 us
Lens mount	CS-mount
Connectors / Power	USB 3.0 micro
Bandwidth	12 MEvents / second
Software	DV-Platform
Power consumption	< 180mA @ 5V DC
Dimensions	H 40 x W 60 x D 25 [mm]
Weight	100g (without lens)
Hardware multi-camera sync	Supported (HiRose Connector)
IMU	6-Axis Built-in
Case	Anodized aluminum, 4 mounting points
Tripod mount	Whitworth ¼" female
APS Frame Shutter	Configurable, Global or Rolling Shutter
CMOS Technology	0.18 um 1P6M MIM CIS
Chip size	8 x 6 [mm]
Pixel size	18.5 x 18.5 [um]
Array size	6.4 x 4.8 [mm]
Fill factor	22 %
Pixel complexity	48 transistors, 2 capacitors, 1 photodiode with micro-lens
Chip voltages	1.8 V and 3.3 V
Chip power consumption	DVS: 10-30mW (activity dependent) APS: 140mW
APS dark signal	18000 e^-/s
APS readout noise	55 e^-

Specifications not guaranteed. All specifications subject to change without notice

Latest change: 15 August 2019

APPENDIX B

Check List

Véronique Quanquin / Sébastien Lengagne Polytech Clermont

Check list – Rapports stage GE4

Les éléments de cette liste doivent être présentés **dans cet ordre** dans le rapport.

Couverture	✓	Logos, titre, prénoms + noms des étudiants, prénoms+noms des tuteurs, mention « rapport de stage » ou « rapport de projet », année, département ; nom de l'école en entier (Polytech Clermont)
Résumé en français et en anglais, mots clés	✓	Résumé (200 mots maximum) + mots clés Abstact + key words Entreprise, problématique, méthode, résultats, analyse des résultats Les 2 résumés sont sur une seule page
Remerciements	✓	Ordre (en gras ce qui est obligatoire) : **tuteur entreprise**, personnels entreprise, **tuteur école**, enseignants école, autres « Mr. » = Mister ; Monsieur = « M. » Attention à l'accord des participes passés
Sommaire	✓	2 niveaux de titres
Table des figures et des tableaux	✓	Toutes les figures et tous les tableaux sont répertoriés. Classement dans l'ordre d'apparition des figures et tableaux dans le texte « Figure 1 : Titre, numéro de page »
Glossaire	✓	Les mots du domaine présentés en ordre **alphabétique** + définition En fin de lexique : une mention indiquant que « tous les mots suivis d'un astérisque sont définis dans le lexique » Dans le texte : un astérisque à chaque mot défini dans le lexique.
Table des abréviations	✓	Obligatoire s'il y a utilisation d'abréviations Présentées par ordre alphabétique
Introduction	✓	Accroche, sujet, problématique, entreprise, enjeux, méthode de travail, annonce du plan
Contexte	✓	Pour évaluer votre travail, il faut clairement définir d'où vous partez et ce que avez à accomplir (de manière chiffrée, par un cahier des charges par exemple)
Résultats	✓	Il faut également définir quels sont vos résultats (quel pourcentage du cahier des charges a été réalisé?).
Figures	✓	Elles sont TOUTES : numérotées, titrées, référencées dans le texte, sourcées
Différences culturelles	✓	Présence d'une partie sur les différences culturelles (en fin de rapport, avant la conclusion)
Conclusion	✓	Rappel de la problématique, rappel synthétique des résultats, distance critique par rapport à ces résultats, ouverture vers un autre sujet ou une autre problématique
Bilan	✓	Prise de recul, analyse des compétences acquises
Bibliographie Sitographie	✓	Application des consignes de présentation des items ne pas confondre bibliographie et sitographie (sites webs d'entreprises par exemple) ATTENTION à bien citer dans le texte **toutes** les références
Table des matières	✓	Tous les niveaux de titres + tous les « objets » du rapport (les tables, les remerciements etc)
Annexes	✓	Table des annexes (« Annexe 1 : TITRE, numéro de page ») les annexes sont numérotées et classées en ordre d'apparition dans le texte, elles sont TOUTES appelées dans le texte
Fiche du tuteur	✓	Présence de la fiche de validation du stage par le tuteur. (c'est une annexe)

Fiche à compléter par l'étudiant et à mettre en Annexe du rapport.

APPENDIX C

Report validation form signed by the tutor

UNIVERSITE CLERMONT-AUVERGNE
POLYTECH CLERMONT

Année universitaire / Academic year: 2022/2023

Attestation de lecture de rapport de stage
Internship report reading certificate

Je soussigné(e), M./Mme
I, the undersigned, Mr or Mrs
Beaussant Samuel

……………………………..…………………………….……………………………

(tuteur d'entreprise), atteste avoir lu et autorise l'envoi au jury le rapport de stage intitulé
(intern's supervisor in the Company) attest to having read and authorizes the sending to the jury of the internship report entitled
"Computer Vision, Reinforcement Learning and Skill Transfer in Robotic: Application to a Dynamic Task"
……………………………..…………………………….……………………………

de l'élève M./Mme
of the student Mr/ Mrs

Antonio Claudio de Sousa dos Santos Filho
……………………………..…………………………….……………………………

Date 30/06/2023
Date 30/06/2023

Signature du tuteur
Company supervisor's signature