Final Master Thesis

MASTER'S DEGREE IN AUTOMATIC CONTROL AND ROBOTICS

Facial Recognition System applied to Multipurpose Assistance robot for Social Human-robot Interaction (MASHI)

MEMORY

Author: Natali Almeida

Advisor: Dr. Cecilio Angulo

Date: September, 2017

ETSEIB

Escuela Técnica Superior

de Ingeniería Industrial de Barcelona

UPC

I would like to thank my God for blessing me to reach as far as I have come because you made this dream come true.

I would like to thank my project tutor, Cecilio Angulo to offer me the opportunity of working together and trust me the development of this work. I would also like to thank my country, Ecuador, for the scholarship offered me to improve my professional and academics skills. Finally,

I would like to thank great friends for the support offered during the master courses and thesis period. To my love Jorge, for its support and unconditional compression, to my parents, Marco and Isabel and my brothers the best friends of my life, for all the trust put over me. And to all my family for your phrases to encourage me to ends this work.

With love Natali.

Abstract

Face recognition is one of the key areas in the field of pattern recognition and artificial intelligence (AI). It has been used in a wide range of applications, such as identity authentication, biometrics, and surveillance.

Image data is high dimensional in the face recognition area, so requires a considerable amount of computing resources and time for recognition. Research effort has been developed in this way, and nowadays many algorithms are available for solving this problem in Computer Vision.

The main goal of this project is to improve the capabilities of the MASHI robot, endowing it for more interaction with humans, and add new functionalities with the components that the robot has.

FISHERFACES, a popular technique for facial recognition is the one chosen to be implemented in our application. This work studies the mathematical fundamentals of this technique to understand how information is processed to perform face recognition. Then, some tests have been performed to check the reliability of the application with several databases of facial images. In this way, it is possible to determine the strengths and weaknesses of the algorithm to be implemented in our robot.

This work introduces an implementation based on Python using the OpenCV library. The characterization of hardware and the description of software is presented. Next, results, limitations, future works, and conclusions over the job development are presented.

ETSEIB

Contents

List of Figures

List of Tables

Acronyms

MASHI Multipurpose Assistance robot for Social Human-robot Interaction

OPENCV Open Computer Vision

HRI Human Robot Interaction

JPEG Joint Photographic Experts Group

RGB Red-Green-Blue

RGBE Red-Green-Blue-Emerland.

YUV Y (brightness) U and V (chrominance color components)

HSV Hue-Saturation-Value

LDA Linear Discriminant Analysis

FLD Fisher Linear Discriminant

PCA Principal Component Analysis

HMI Human Machine Interface

SHRI Social Human Robot Interactions

LUT Look-Up Table

ROI Region Of Interest.

UI User Interface

AI Artificial Intelligence

AP Access Point

CPU Central Processing Unit

DOF Degrees of Freedom

IP Internet Protocol

LCD Liquid Crystal Display

OS Operating System

PC Personal Computer

PWM Pulse Width Modulation

RPi Raspberry Pi

USB Universal Serial Bus

FPS Frame per Second

TOPIO TOSY Ping Pong Playing Robot

TP True Positive

TN True Negative

FP False Positive

FN False Negative

Chapter 1

Introduction

One of the important and complex activities of a social robot is the implementation of resources of communication that allow the interaction between the human and the machine. A social robot should be endowed with detection skills and capture the basic indications from human behavior understanding; it must be able to make interactive exchanges with its environment in a natural way.

The goal of computer vision is to understand the story described in a picture. As humans, this is quite simple; but for computers, the task is extremely difficult. Currently, robotic projects aimed at the construction of social robots focus on the ability of the robot to recognize faces and associate them with people. The area of computer vision involves both, the processing of digital images and the development of techniques that enable the association of a face with that of a known individual if it is in the robot's database. This work proposes the development of a software tool to identify a face, extract its characteristics and store them in a database.

1.1 Motivation

At present, the Universitat Politècnica de Catalunya (UPC) owns an educational robotic platform called MASHI. Students from the UPC have rebuilt this robot and modified its original design. For instance, the processing unit CPU in the original idea was a laptop and it was switched to a Raspberry Pi 3 with the aim of reducing energy consumption, reduce costs, work with open source code, and minimize the physical space at the base of the robot.

The main motivation for this work is to increase the social interaction capacity of the MASHI robot, through improvements in the hardware and implementation of new applications in the software.

The necessary work in the hardware was completed so that the MASHI platform is fully enabled. It can work independently and can be used by more students in the future for the implementation of different projects.

On the software side, the facial recognition program was developed over the Raspberry Pi 3 using the programming language Python. This work enables the use of the frontal camera of the robot, which performs the training of the robot to learn faces and then identify the people who appear in front of MASHI.

The development of different applications on MASHI allows students to apply coding strategies in projects and use the different resources learned in the career to expand the capabilities of the robot.

1.2 Objectives

The overall objective of this thesis is *Develop a human-robot interaction system, based on computer vision, for the social robot platform MASHI and the study of the needs of a recognition system and learning-based on vision. The implemented prototype must recognize and identify people inside a building. Furthermore, the tests must be performed by the implemented algorithm, analyzing the different extrinsic factors of the environment conditions (illumination, distance), and how this affects the perception and interpretation of the MASHI robot.*

1.3 Document structure

The document is structured in the following chapters:

Chapter 1. Introduction. First chapter contains a brief introduction about the interaction between humans and robots, as well as the motivation of the project and the objective to be achieved with the development of the work presented.

Chapter 2. State of Art. In the second chapter, the state-of-the-art is presented. The notions that affect the project, needed to understand the concepts used in the development of the work, begins with a review of previous and parallel projects about socially interactive robots. Their configuration, the consequences of the interaction with humans, as well as the techniques of computer vision and image processing focused on facial recognition are also discussed.

Chapter 3. Methodology. The third chapter performs a study of the methodology implemented, including the tools and software used. The vision algorithm developed for facial recognition over the MASHI robot is described, alongside a section dedicated to the MASHI training, and another section with the implemented technique of identification of people using Fisher Recognizer.

Chapter 4. Implementation and Testing The fourth chapter shows the test results obtained with the MASHI robot with different external factors, such as training distance, luminosity; and changed internal factors, like the number of photos that are used for learning the face.

Chapter 5. Cost. Chapters five and six cover the breakdown of material cost and environmental impact respectively.

Chapter 6. Environmental Impact.

Chapter 7. Conclusions and Future Work. Finally, the conclusions obtained from the accomplishment of this work are explained, as well as future developments that can derive from this work.

Chapter 2

State of the Art

2.1 Human robot interaction

Imagining a world with robots in our daily lives seemed unachievable a few years ago. Nowadays, with the swift technological advances, robots are getting familiarized with our daily activities and becoming a part of our society by holding positions, such as medical assistants, social workers, industrial workers, educational robots, pets, etc. A considerable progress is being made in this direction; many laboratories around the world carry out interdisciplinary research to develop robots that are useful and safe for everyday life. Moreover, they develop research around the interaction [1] between robots and people to guarantee efficient, ethical and responsible use. Some areas in which robotics has been put to used is shown in Figure 2.1.

Figure 2.1: Robots today.

2.2 Socially interactive robots

With socially interactive robots, the important goal to be achieved is that the communication with the environment seems natural [2]. The design development of robots that allows the implementation of interfaces for efficient communication between humans and machines is a great challenge.

The use of devices such as keyboards, screens, mouses and joysticks to control a robot remotely cannot be considered social, since the robot is not able to make decisions by itself, the human only manages it. A complete autonomy does not mean that a robot is to be considered socially interactive. Partial autonomy is also allowed, this skill uses similar communication mechanisms to those used by humans, such as visual communication, touch, gestures, voice, etc. Figure 2.2 and Figure 2.3 illustrate examples of different levels of autonomy.

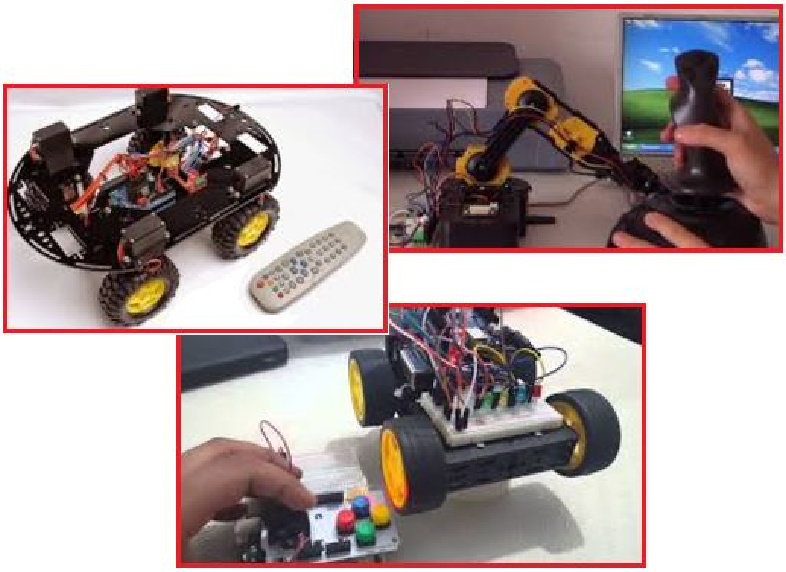


Figure 2.2: Robots controlled.

In this context, autonomy stands for the independence that the robot shows from the control performed by humans. Therefore, an autonomous robot must be able [4]:

- To store and learn through the information of the environment.
- To react in an environment with continuous changes.
- To act without assistance in considerable time periods.
- To avoid dangerous situations for the people as well as for itself.

Figure 2.3: Social interactive robots.

2.2.1 Social robotics

The earliest research in the area of AI was in 1943, when mathematician Walter Pitts and neurophysiologist Warren McCulloch presented the first research work on the way in which neurons interact in the brain; and proposed alongside the construction of machines that mimic the functioning of human neural networks [3]. In this way, they both founded the basis of artificial neural networks. Social robotics had its beginnings in 1948 when William Gray Walter built the robots known as Elmer and Elsie (see Figure 2.4).

He showed that with a small number of connections these robots could develop relatively complex behaviors as overcome obstacles, return to their burrow, react to light or touch stimuli, and recharge their batteries before they were depleted.

During the second half of the twentieth century successive generations of industrial robots were programmed to do repetitive tasks. At the same time, some ideas to promote projects of humanoid robots were developed; these robots were modified in shape to get away from the design of industrial robots and were given the skills to interact with people in health-care, urban and service environments.

2.2.2 Characteristics of socially interactive robots

The mains skills of socially interactive robots in domains where robots must exhibit peer-to-peer interaction skills are the following [5]:

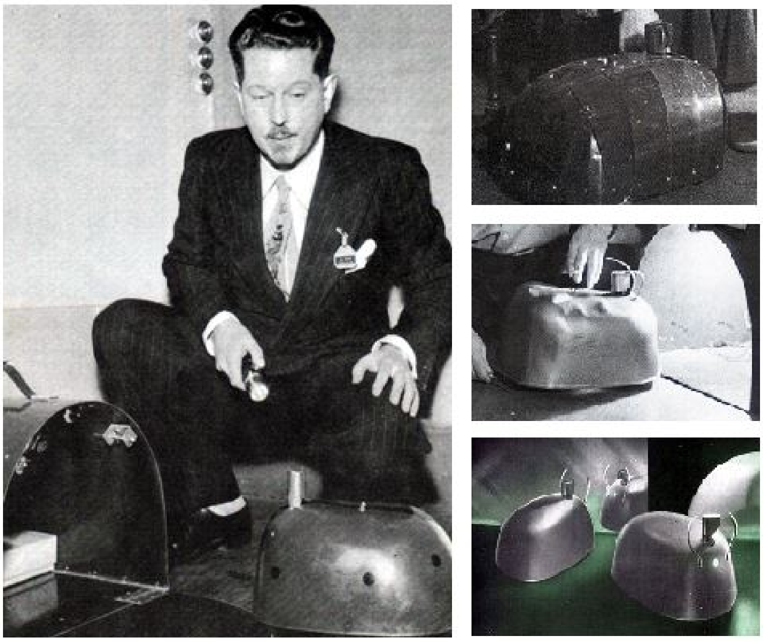


Figure 2.4: First social robots.

Establish and maintain social relationships. Taken into account that socialization with people is a difficult topic, consider the additional difficulty where robots and humans do not share a common language nor perceive the world in the same way.

Perceive and/or express emotions. A robot must be able to recognize different moods, like joy, sadness, anger, etc.

Communicate with high-level dialogue. Robots must be able to follow the movements of the interlocutor and must be able to recognize and interpret human speech, which includes discourse, discrete commands, and natural language.

Learn/recognize models of other agents. Robots must be secure, capable of perceiving and manipulating deformable perceptual-tolerant objects and inaccurate actions in dynamic environments and must be endowed with a great capacity for learning and adaptability to non-predefined environments.

Use natural cues (voice, gaze, gestures, etc). It is known that humans feel more confident when robots have a certain degree of expressive capacity. That is, the design must have a face, with primary facial components such as eyes, eyelids, eyebrows, cheeks, lips, and jaw.

Exhibit distinctive personality and character. People are more willing to interact and establish a relationship if the robot that can provide useful feedback in every situation.

May learn/develop social competencies. For example, this entails helping older people or people with reduced abilities who experience difficulties with eating, taking medication, accessing places, writing and texting, taking photographs, collecting information, or general support.

2.3 Related works

Currently, there are several projects that focus on the development of social robots, which present many differences between them regarding communication, recognition, cooperation, performance, morphology and other characteristics. This section introduces related works that help define the different interaction goals for these types of projects. Also, the information presented is intended to clarify the mode of operation of socially interactive robots in which computer vision algorithms have been applied.

2.3.1 NAO robot and social interaction

The company Aldebaran Robotics developed the humanoid robot called NAO [6]. The artificial vision of NAO uses some techniques such as: image segmentation, object detection, depth estimation, stereo-vision, integration between kinematic model, and robot perception model [7].
NAO was created as a service robot to help children, the elderly, firemen, police, etc. The robot can walk, talk, listen, dance, sing and interact with the environment around it.
Considering a research performed by the GREC-UPC research group in the Margalló Elementary School, the social interaction between children and the NAO platform is very pleasing [9]. Children are comfortable with a robot that has human morphological characteristics because they can engage in amusing activities with the robot. However, it should be emphasized that not all the expectations of the children were fulfilled, since they wanted the robot to have more skills such as thinking, recognizing, or writing. Even then, the conclusion was that everyone had a fun time, and the children wanted to interact with the robot again. Social interaction between NAO and children can be seen in Figure 2.5 (a).

2.3.2 TOPIO robot and social interaction

The company TOSY developed the humanoid robot TOPIO to play table tennis against human beings. Since 2005, three robot prototypes have been developed (See Figure 2.5 (b)). The following Table 2.1 shows the scope in artificial vision.

In the week-long Robot Exhibition in China (IREX) 2009 [13], TOPIO's debut was impressive because it showed special skills like accuracy and fast speed in its movements.

Model	Topio 1.0	Topio 2.0	Topio 3.0
Artificial Vision	Recognize the rotation and trajectory of the ball	Estimate the ball trajectory and update successively based on the image processing results	96% for ball detection. + adaptive path planning algorithm is designed for robot motion

Table 2.1: Topio artificial vision.

Nonetheless people complain that the robot cannot perform complicated ping-pong techniques such as slicing and curving.

2.3.3 MASHI Platform

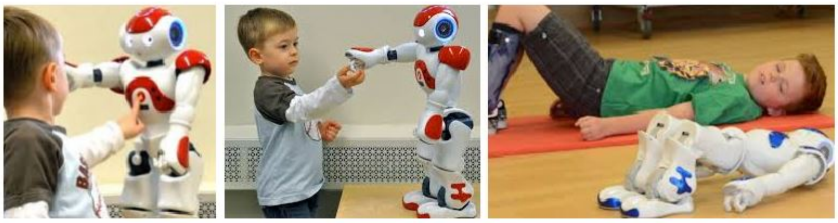
The reconstruction work of the MASHI platform and the explanation of the morphology of the robot was developed in two projects [11], [10].
In [11] a path-finding function was developed in the robot and concluded that the algorithm works correctly according to the simulation data. Future work involves improving the sensing and mobility capabilities. Some topics were developed in the maintenance work of MASHI [12].
In [10] the change of the control system of the MASHI robot was proposed. Here, the laptop was changed to a Raspberry Pi 3 (RPi 3) with the purpose of having a lighter and more economic control system that maintained the general functionalities of the platform. Henceforth, the obtained results were satisfactory considering that the functions of movement and connectivity of the MASHI were supported by the RPi3

2.3.4 MASHI robot and social interaction

The authors in [14] summarize the description about social human robot interactions (SHRI) and propose an experiment based on an observational method to evaluate the space, F-formation, and the proxemics behavior, in an environment that humans and robots are sharing. For this experiment the MASHI, an experimental robotic platform for social human-robot interaction research was used. The result of this research showed that most of the people that interact with the machine were children and young people, who are mainly searching for entertainment or to have a dialog, while keeping a personal distance. Figure 2.5 (c). shows the MASHI social interaction.

2.3.5 Face recognition algorithm in social robots

Human Robot Interaction (HRI) is the requirement that allows the robots to share a social environment with human beings. One of the main functionalities of HRI is face detection and recognition, because it offers a friendly interface between humans and robots. In that way, the author in [15] has proposed a solution for face tracking and detection via the use of AdaBoost and PCA algorithms that are executed on an embedded processor that makes part of a humanoid robot named Philos [16]. The results obtained were evaluated considering different luminosity and verifying the level of reliability of the algorithm, to assess the recommendation of this solution for similar applications in HRI solutions.



(a) NAO's social interaction with children.

(b) TOPIO robot exhibition.

(c) MASHI social interaction.

Figure 2.5: NAO robot, TOPIO robot and MASHI robot.

2.4 MASHI Platform

MASHI is a Multipurpose Assistance humanoid robot designed to fulfill a companionship role in order to study Human-Robot Social Interaction. This robot was created by researchers of the UPC in the year 2014. Currently, there are two prototypes one in Ecuador and another in Spain, where the difference between the two models lies in the controller used to manage the different activities that the robot can perform.

2.4.1 MASHI-Ecuador

The prototype of MASHI in Ecuador is controlled by a computer and can perform the following activities: hold a conversation, make selfies, shake the hand of whoever approaches it, teleoperated movements and change the facial gestures. Figure 2.6. shows the prototype of MASHI in Ecuador and its designer Dennys Paillacho.

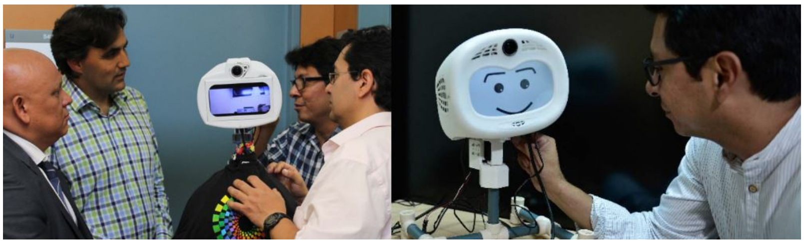


Figure 2.6: MASHI Robot in Ecuador.

2.4.2 MASHI-Spain

MASHI-Spain is a copy of the MASHI-Ecuador prototype and was rebuilt in Barcelona by students Xavier Rodríguez and Joaquín Cortés in 2016. The aim of this project was to create a framework that allows UPC students to implement continuous improvements on the platform through the development of new projects. The presented work was be realized on this robot. Figure 2.7 shows MASHI at the UPC.

Figure 2.7: MASHI Robot in UPC.

2.5 Work developed over MASHI-UPC

Unlike the initial model, the MASHI robot of the UPC uses a RPi 3 as a processor instead of a computer. This reduces the cost of the materials involved in the reconstruction of MASHI. To develop the social interaction analysis of MASHI using a RPI 3 and considering the current state of this platform [12], the following work was developed before the implementation of the code.

Quadrature encoders replacement: The original motors mounted on the MASHI's base contained two Parallax position controllers part number 27906. After doing the maintenance, it was found that the serial communication (UART) bus did not work correctly, hence the replacement of these sensors by the encoders part number 29321 of the same manufacturer, was performed. In addition, the code to create the data that the old sensor sends through the UART was developed, in this way the measurements to control the movement of the platform can be used for different projects in the future. Figure 2.8 shows this work

(a) Screen power cable. (b) Motor's power cable.

Figure 2.8: MASHI's cables to supply the screen and motors.

Wiring and labeling: To ensure the correct connection between the elements of MASHI and considering the current state of the encoders and cables of the platform, the cables were changed and labeled to help other students in future projects. Figure 2.9 shows this work.

(a) Screen power cable. (b) Motor's power cable. (c) Current state.

Figure 2.9: MASHI's cables to supply the screen and motors.

Selfie skill: When the maintenance of the MASHI robot was carried out, it was not able to take selfies and send the photographs to MASHI's twitter account. To allow this functionality, parameters in the Java encoding on the RPi3 were reviewed and changed. Figure 2.10 shows this work and Figure 2.11 shows some pictures of the MASHI's twitter profile in the Maker Faire at Barcelona 2017.

Figure 2.10: MASHI taking a selfie.

Figure 2.11: MASHI taking a selfie.

Mechanical Maintenance: Mechanical tasks were performed to enable all the motors that allow the movement of MASHI. As can be seen in Figure 2.12 the current MASHI in the laboratory is able to move the arms and the head and also change facial gestures.

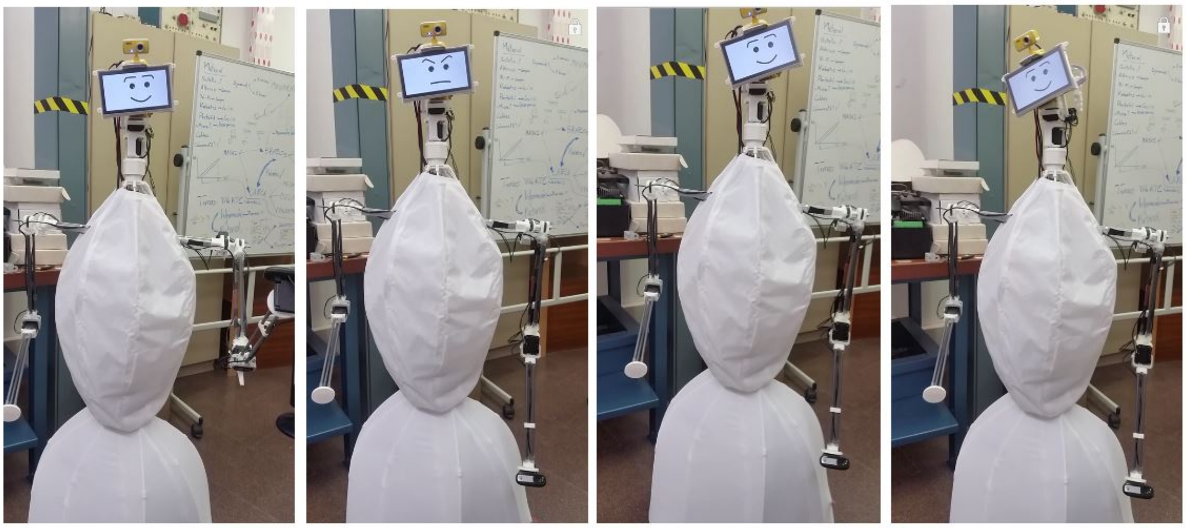


Figure 2.12: MASHI current state.

Chapter 3

Methodology

3.1 System description

As previously mentioned, this work consists on the development of a facial recognition system for the humanoid robot MASHI through the use of artificial vision. The robot will be able to recognize the people who are in the database and frame the face with the name of the person who is in front of MASHI. The system must be able to:

- Store the faces and save the data with the corresponding name.
- The images of the faces stored in the system will have to be differentiated to be identified.

This project was developed with the frontal camera of MASHI. This sensor did not have a function associated in previous projects and thus the current functions of the robot are maintained; in other words: the robot can move, take selfies and also perform facial recognition.

3.2 Technologies and tools

To carry out the described functionalities. The system will consist of the following:

Software:

- The MASHI operating system is Rasbian GNU / Linux 8.0 (Jessie).
- Java Version 1.7.0-121 + JavaScript with Node.js V0.10.29 + Geany Software for HMI
- Python 2.7 has been used for the development of artificial vision. Along with the libraries of OpenCV version 2.4.9 and Numpy 1.8.2 (Dependency of OpenCV), for Python according to the operating system.

Hardware:

- The MASHI robot
- Raspberry Pi 3
- Arduino Mega 2560

3.2.1 OPENCV

The OpenCV library (The Open Computer Vision Library), was born under BSD license [17]. Today, it has a community of more than 47,000 people and more than 7 million downloads. It is a commercially used library, companies like Google, Yahoo, Microsoft, Intel, IBM, Sony, Honda, Toyota, etc. use it in their applications.

OpenCV provides an infrastructure for the development of real-time computer vision applications such as object identification, data structures, structural analysis, 3D reconstruction, augmented reality, segmentation and object recognition, etc.
OpenCV is written in C ++, has interfaces in C ++, C, Python, Java, and MATLAB and works on Windows, Linux, Android and Mac OS.

3.2.1.1 OpenCV library structure

OpenCV has a modular structure. Figure 3.1 shows the main modules.

Core. It is a basic module. It includes the basic data structures and the basic functions of image processing.

Highgui. This module provides UI, image, video codecs, and the ability to capture images and video. It allows writing/reading images in numerous formats (BMP, JPEG, TIFF, PxM, SunRaster, etc.).

Imgproc. This module includes basic algorithms of image processing, including filtering of images, transforming of images, etc.

Video. This module serves for video analysis and includes object tracking algorithms.

Objdetect. This module includes Object Recognition and Recognition algorithms for standard objects.

3.2.2 MASHI robot

The MASHI robot is a platform whose skeleton is composed almost entirely of plastic PLA 3D. The MASHI has 8 DOF, distributed as follows:

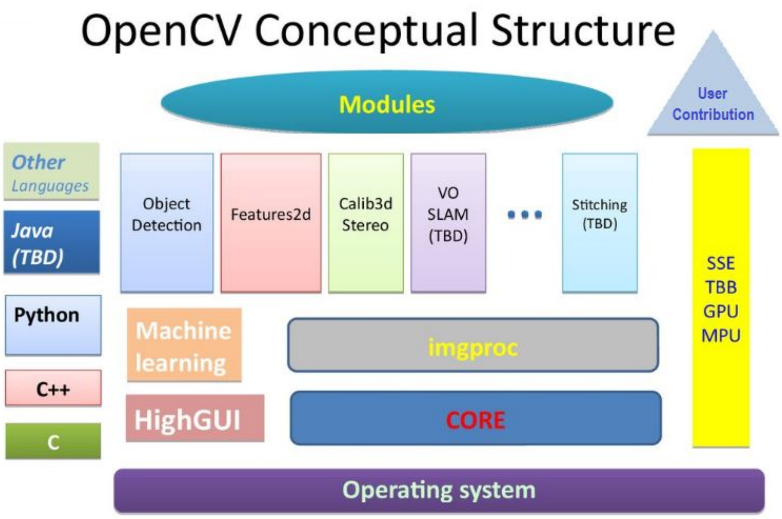
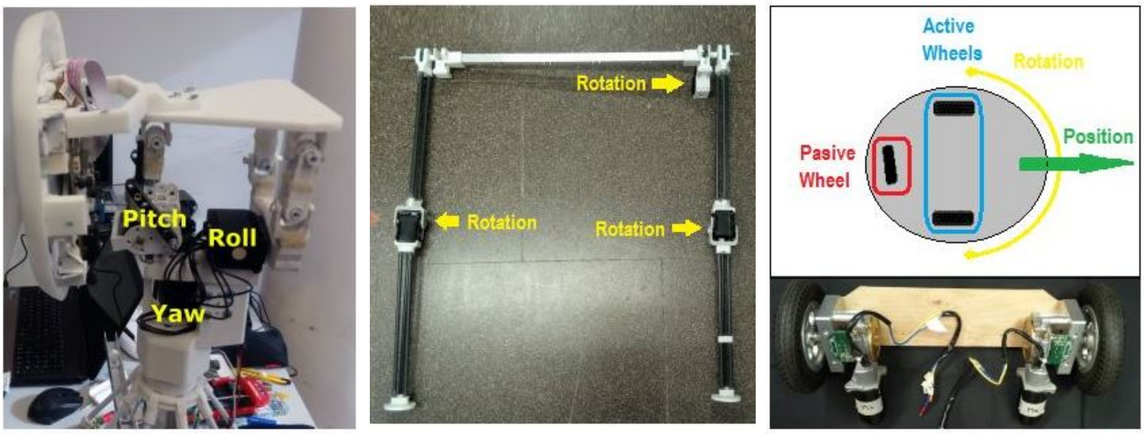


Figure 3.1: OpenCV modular structure.

- 3 DOF on the head: Yaw, pitch, and roll.
- 2 DOF on the left arm: Corresponds to the rotation of the shoulder and the rotation of the elbow.
- 1 DOF on the right arm: Corresponds to the rotation of the elbow.
- 2 DOF at the base: Corresponds to the position and orientation of the robot.

In Figure 3.2 the DOFs of the MASHI robot can be observed.



(a) Head (b) Arms (c) Base

Figure 3.2: MASHI DOF.

The MASHI robot is endowed with two cameras: one on the head and one on the left arm with a resolution of 640×480 at 30 frames per second (fps). The location of the two cameras are shown in Figure 3.3.

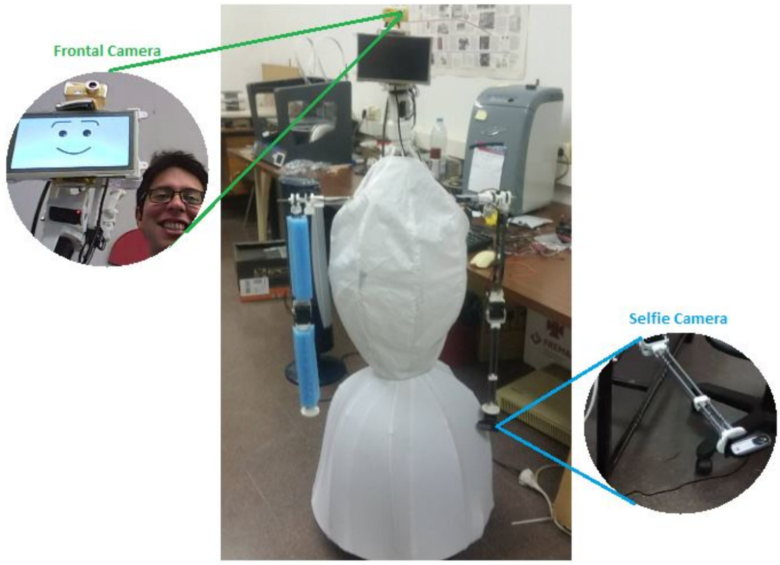
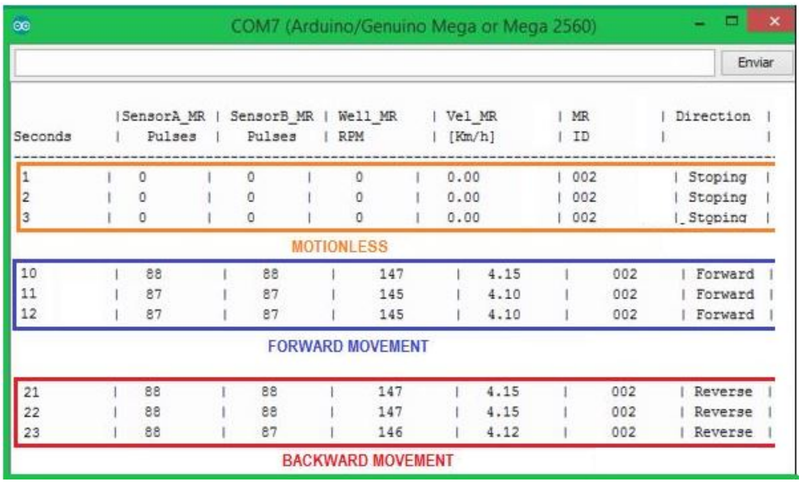


Figure 3.3: Cameras in MASHI.

Finally, MASHI has two encoders of quadrature in each of the wheels in the base. These encoders send two overlapping signals that indicate if the robot moves or not, by calculating the speed and the distance traveled by MASHI. All this data is processed in the Arduino Mega 2560 and the results for each case is summarized in Figure 3.4.



COM7 (Arduino/Genuino Mega or Mega 2560)

Enviar

Seconds	SensorA_MR Pulses	SensorB_MR Pulses	Well_MR RPM	Vel_MR [Km/h]	MR ID	Direction
1	0	0	0	0.00	002	Stoping
2	0	0	0	0.00	002	Stoping
3	0	0	0	0.00	002	Stoping

MOTIONLESS

Seconds	SensorA_MR Pulses	SensorB_MR Pulses	Well_MR RPM	Vel_MR [Km/h]	MR ID	Direction
10	88	88	147	4.15	002	Forward
11	87	87	145	4.10	002	Forward
12	87	87	145	4.10	002	Forward

FORWARD MOVEMENT

Seconds	SensorA_MR Pulses	SensorB_MR Pulses	Well_MR RPM	Vel_MR [Km/h]	MR ID	Direction
21	88	88	147	4.15	002	Reverse
22	88	88	147	4.15	002	Reverse
23	88	87	146	4.12	002	Reverse

BACKWARD MOVEMENT

Figure 3.4: Quadrature encoders data.

3.2.3 Raspberry Pi 3

The Raspberry Pi 3 is the third version of the embedded platform of the same name. It is a complete low-power computer on a single board. This brings several improvements with respect to its previous model, of which the following stand out:

- Quad Core Broadcom BCM2837 64-bit ARMv8 processor. This processor has speeds of up to 1.2GHz compared to the previous 900MHz on the Pi 2.
- 1GB RAM Memory.
- The addition of a built-in BCM43143 WiFi chip, allowing the Pi 3 to go wireless without additional peripherals.
- The power source has been upgraded to 2.5A instead of 2A, allowing the Pi to power even more powerful devices over USB ports.
- The Pi 3 has four built-in USB ports for connectivity to a mouse, keyboard, or anything with a USB connection.

The following diagram in Figure 3.5 shows the connection of the Raspberry Pi in the MASHI robot.

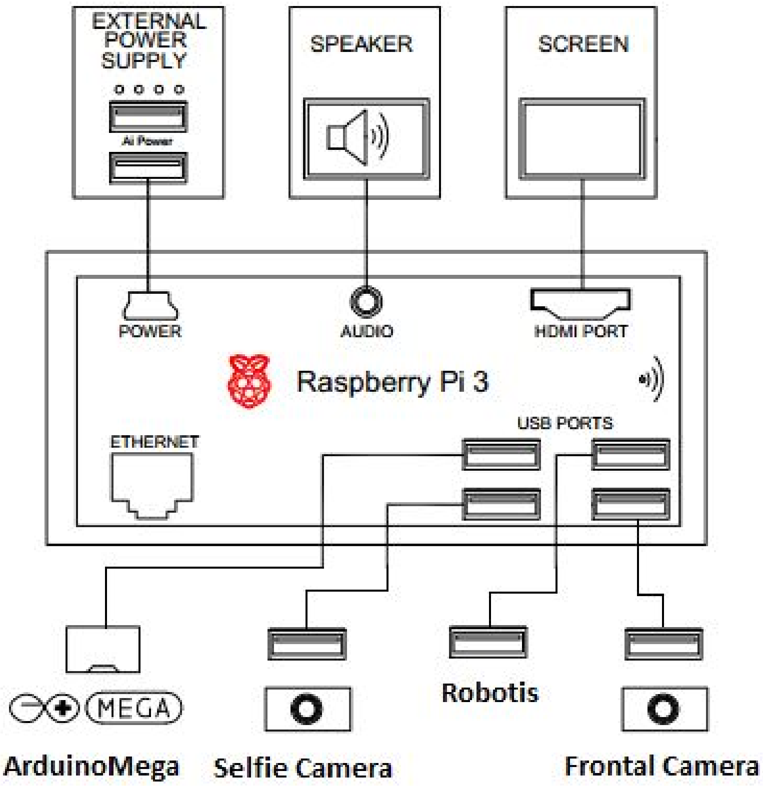


Figure 3.5: Raspebrry Pi 3 connection diagram.

3.3 Facial recognition algorithm

The facial recognition algorithm has two defined phases:

- Training phase: In this phase the database is prepared to perform the recognition. Several images of people's faces are used to train the system. In this phase, the extraction of characteristics of each person is carried out and stored for later comparison.

- Phase of recognition or test: In this phase, the images of an unknown subject are taken, and the extraction of characteristics is carried out with the same process of the first phase, to compare these characteristics with the characteristics of the database.

The process that the algorithm follows is described below [18]:

1. **Acquisition of the image:** Depending on the system used, it can be a still an image, a video frame, a three-dimensional image, etc.

2. **Face detection:** Uses object detection algorithms that detect whether there is a face in the image. Provides the location and size of the face.

3. **Image processing:** When a face is determined on the image, the normalization process is performed. Face components such as size, pose and illumination are located. To normalize face images, different rules can be followed, such as the position of the nose, the distance between the pupils of the eyes, and the size of the lips.
   To optimize the system's performance, processes such as reducing the image size, converting the image to gray scale, or using a low-pass filter are used if the image resolution is too high.

4. **Extraction of the characteristics:** After the image is processed, the characteristic vectors or coefficients of the image will be calculated depending on the technique used.

5. **Recognition:** Finally, the extracted feature vector is compared with the feature vectors extracted from the faces database. If it finds one with a high percentage of similarity, it returns the identity of the face; if not, it indicates that it is an unknown face.

The diagram corresponding to the previously detailed flow diagram is illustrated in Figure 3.6.

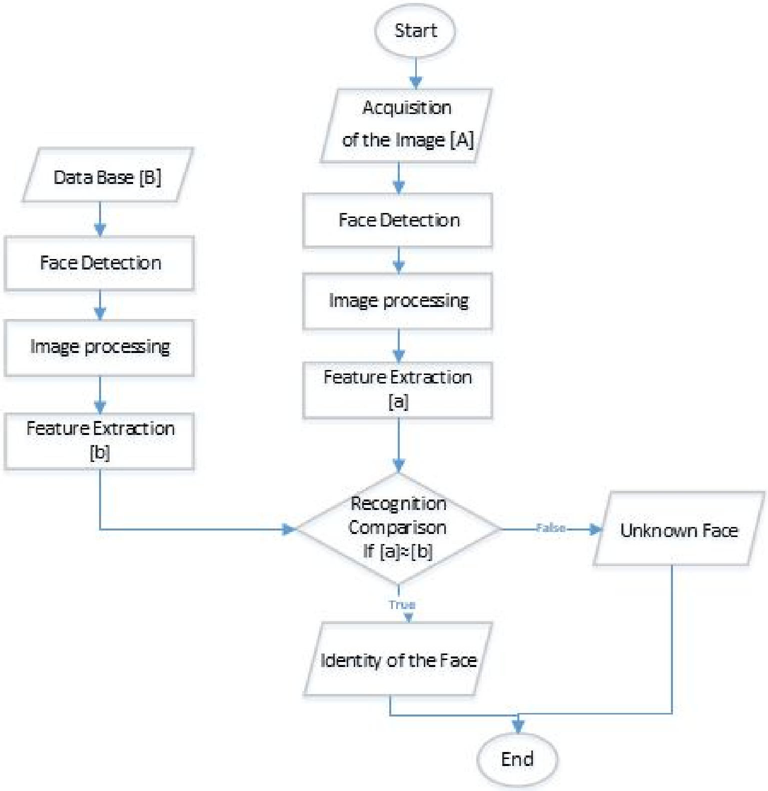


Figure 3.6: Facial recognition flowchart.

3.3.1 Face detection

The first thing to do in a facial recognition system is to detect a face in an image. To achieve the face detection, the algorithm must contemplate many facial structures and the relative position of the different elements that make up the face: the structure of the eyes and nose, the distance between the eyes and nose, the size and shape of the face, etc. The identification process is shown in Figure 3.7. One of the most used algorithms to detect objects in real-time is the fast object detector of Viola-Jones.

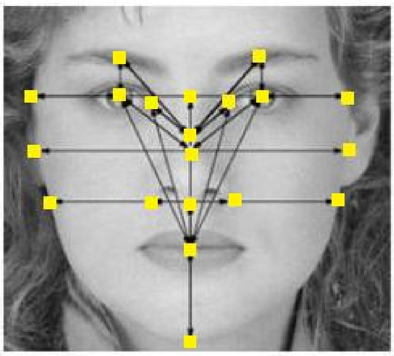


Figure 3.7: Facial structures.

3.3.1.1 Viola-Jones algorithm

The Viola-Jones method is one of the most used processes today [19] because it allows segmenting multiple faces in an image with low processing times. Viola and Jones based the algorithm on simple features instead of pixels. To do this it uses a Haar base for the extraction of characteristics and Adaboost for the selection and classification of characteristics. The flowchart is divided into three stages as shown in Figure 3.8:

1. **Transformation:** Generates a new image called "integral image".
2. **Feature extraction:** Using Haar-based filters.
3. **Construction of cascade classifiers:** Using boosting.

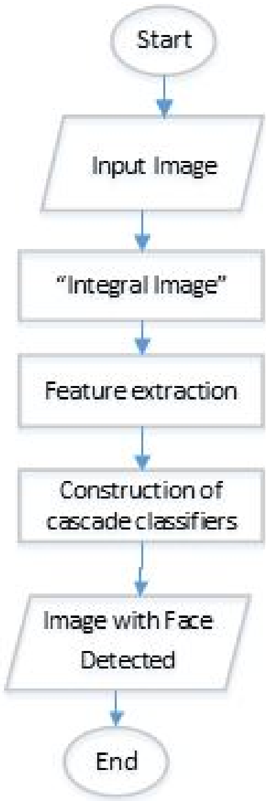


Figure 3.8: Viola-Jones flowchart.

Integral image

An integral image is a cumulative image of the original image that allows to calculate the sum of the values of the pixels in any rectangular area of a given image [20]. The integral of an image with respect to a point (x, y) consists of the sum of the pixels above and to the left of said points, (x, y) included, as shown in the Figure 3.9.

$$ii(x,y) = \sum_{x' \le x, y' \le y} i(x',y')$$

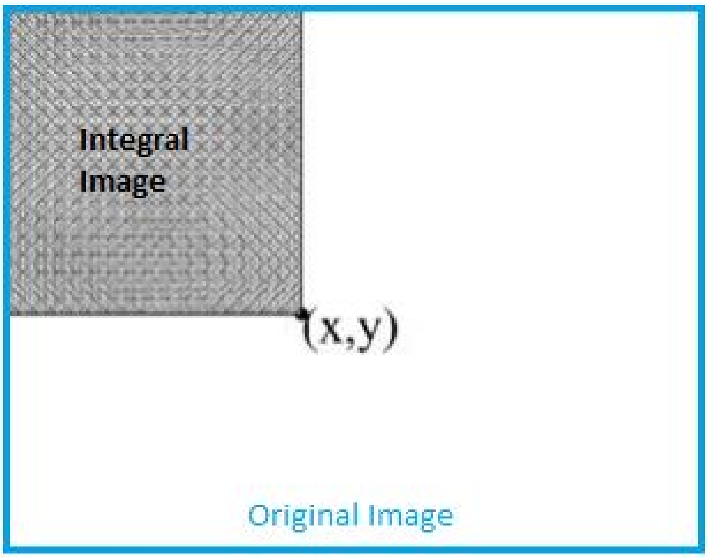


Figure 3.9: Pixels used to represent the integral image of a pixel (x, y).

Feature extraction

The extraction of characteristics is carried out by applying image filters based on Haar over the integral image. These characteristics are calculated as the difference of the sum of the pixels of two or more adjacent rectangular zones [19]. This algorithm uses three types of Haar characteristics:

1. **The two-rectangles feature**, which is the difference between the sum of the pixels of two Rectangles. These rectangles have the same shape and are adjacent vertical or horizontally.

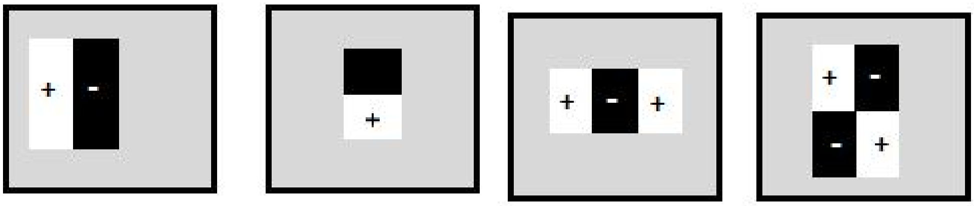
2. **The three-rectangle feature**, which computes the sum of the pixels within two outer rectangles, in turn, subtracted from the sum of a third inner rectangle.

3. **The four-rectangle feature**, which computes the difference of paired rectangles in diagonal.

The Haar features explained above are illustrated in Figure 3.10.
An example of a feature extraction in a picture is shown in Figure 3.11 .

Construction of cascade classifiers

Boosting is a classification method that uses basic classifiers (Adaboost) to form a single classifier more complex and precise [20]. This algorithm adds simple classifiers one after



(a) Two rectangles Feature (b) Three rectangles Feature (c) Four rectangles Feature

Figure 3.10: Haar features.

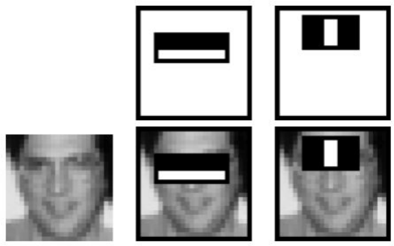


Figure 3.11: Feature extraction.

another, each one with a slightly higher accuracy than a random classification and combines them to get a much more accurate classifier. This classifier is shown in Figure 3.12.

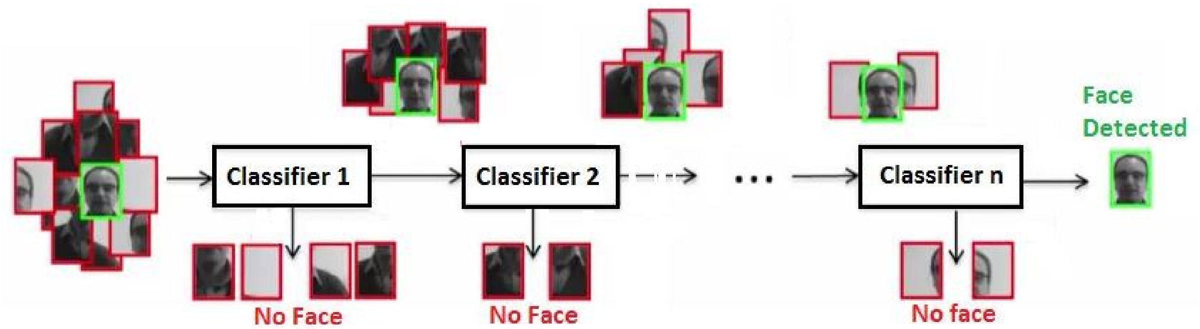


Figure 3.12: Cascade classification.

3.3.2 Recognition: Fisherfaces

Fisherfaces is a technique based on fixed images. This method was proposed by Belhumeur in which PCA and the Fisher Linear Discriminant (FLD) or LDA are used [21].

3.3.2.1 Principal Component Analysis (PCA)

The Principal Component Analysis technique has two phases [22]:

- **Training Phase:** Using PCA, creates the eigenspace (faction space) from the facial training images. The factions space is the matrix formed by the eigenvectors. These vectors contain information on the variation of the gray values of each pixel of the set of images used.

  To form the space of factions (eigenfaces), the first vectors of the matrix are used because they contain the most important information of the space, in this way a significant reduction of the managed information is achieved. An example of eigenfaces representation can be seen in Figure 3.13. To conclude this phase, the images that were used when performing the PCA are projected in the eigenface. The projection characterizes the facial image of an individual as the sum of the different weights of the image space.

- **Classification Phase:** An unknown facial image is projected against the eigenspace using the Euclidean distance, looking for the projected facial image more similar to the unknown.

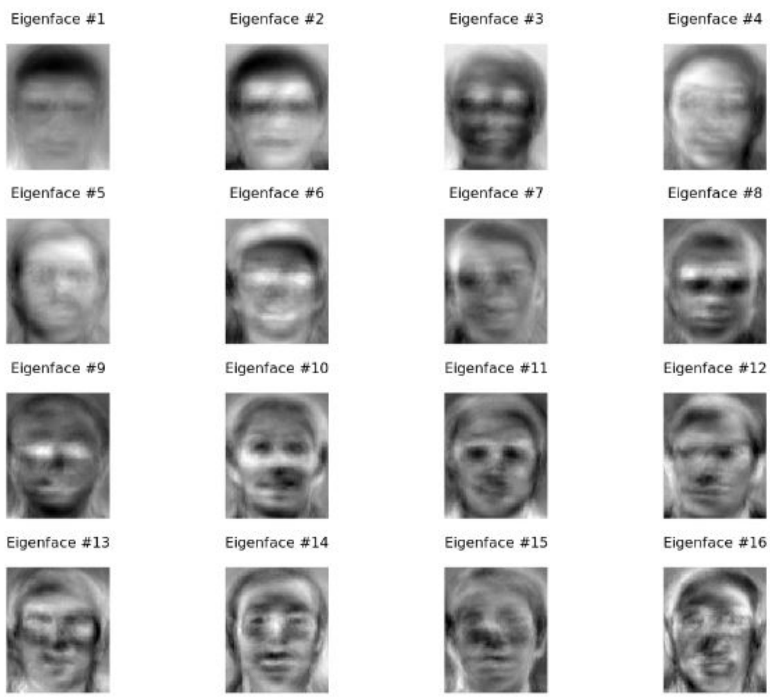
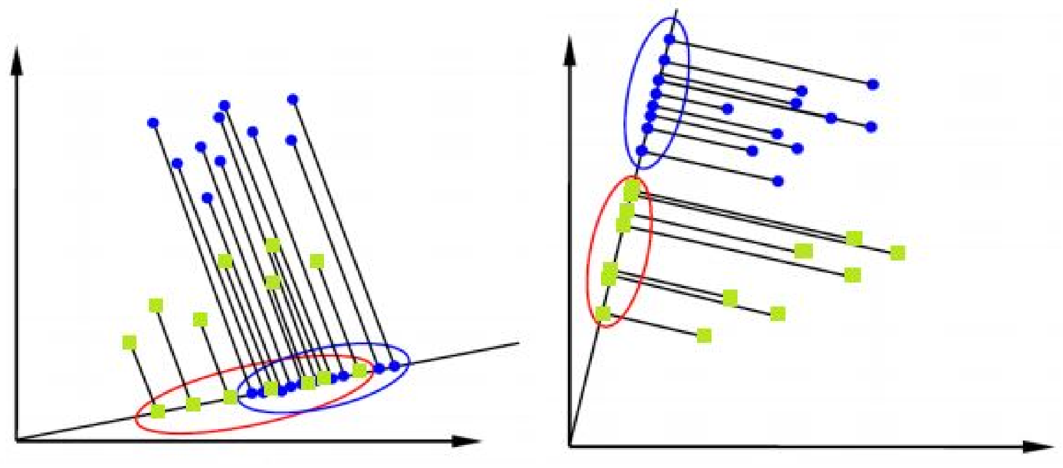


Figure 3.13: Eigenfaces representation.

3.3.2.2 Linear Discriminant Analysis (LDA)

The Linear Discriminant Analysis technique maximizes the variance of samples between classes (different people) and minimizes it between samples of the same class (of the same person).The idea is simple: same classes should cluster tightly together, while different classes are as far away as possible from each other in the lower-dimensional representation [23].
The recognition phase is the same as in the PCA: first, the image is projected into the space formed by the training images, and then searches for the more similar image through Euclidean distance. The difference between both techniques lies in the way the space is calculated. To calculate the eigenspace, two covariance matrices are searched, one of "inter-class" corresponding to the different images of the same person, and other of "extra-class" corresponding to the images of different people. Two different ways of projecting the same points on a line, one seeking to minimize the distance between points of the same class, and another looking to maximize the distance between points of different class are illustrated in Figure 3.14.



(a) Minimizing the distance.

(b) Maximizing the distance.

Figure 3.14: Points projected in a line.

The relationship between the two matrices is known as the LDA subspace, composed of the known eigenvectors of the PCA technique. These eigenvectors are the Fisherfaces that give the name to this technique. An example of Fisherface representation is shown in Figure 3.15.

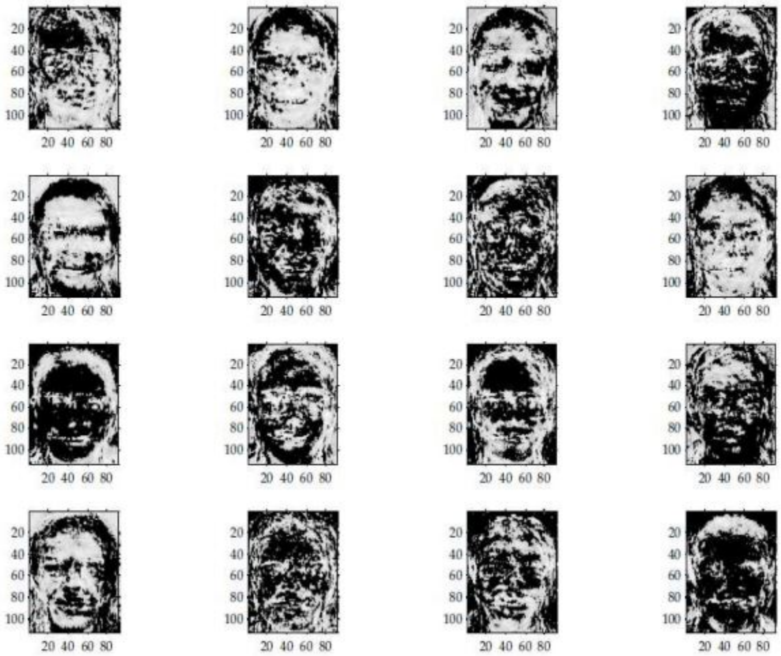


Figure 3.15: Fisherfaces representation.

Chapter 4

Implementation and Testing

4.1 Project Overview

In the present project, an application for facial recognition was implemented in Python over the RPi3 of the MASHI robot. To achieve this, the library OpenCV that supports image processing has been used.

The algorithm implemented for the facial recognition problem is divided into two different and independent modules:

1. Acquisition of the subject's facial image to identify and the processing this image.

2. Recognition technique to be used: FLD (Fisher Linear Discriminant).

4.1.1 Acquisition of facial image

In the training phase of MASHI the user needs run the "capture.py" application + "name" (Figure 4.1). This application open the "Face Training" window and develops the following:

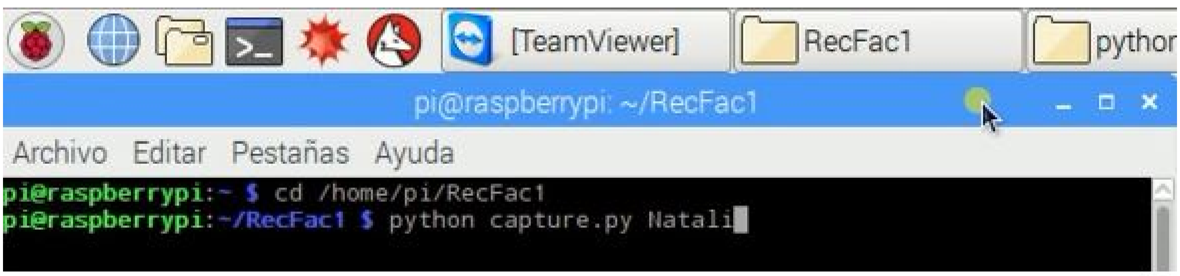


Figure 4.1: capture.py application.

- Open the front webcam of the MASHI robot, take photos and apply a classifier to find facial images (see Figure 4.2). The algorithm allows to save a set of the face images in a folder with the name of the subject.

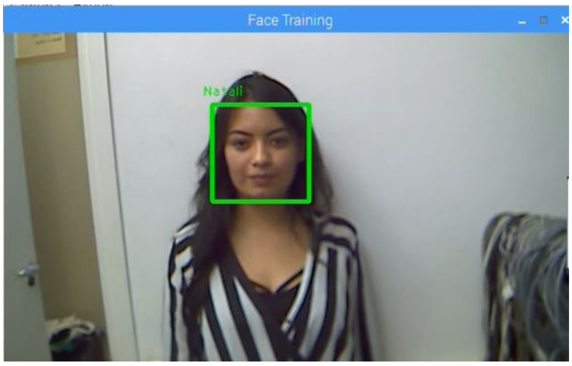


Figure 4.2: Face detection with the application.

- Firstly crop the images and process them so that they are grayscale and have the same size in pixels. The software also equalizes the histogram of images to reduce the impact of ambient light variations when capturing images (see Figure 4.3).

Figure 4.3: Faces processed with the application.

- Once the application has determined the faces of the people, save the images in the database of the robot in a folder with the name of the person that MASHI should recognize (see Figure 4.4).

Figure 4.4: Data base with the application.

The OpenCV library comes with functions and classes that allow the development of the above-described algorithm. A flowchart illustrating the process for face recognition is shown in Figure 4.5.

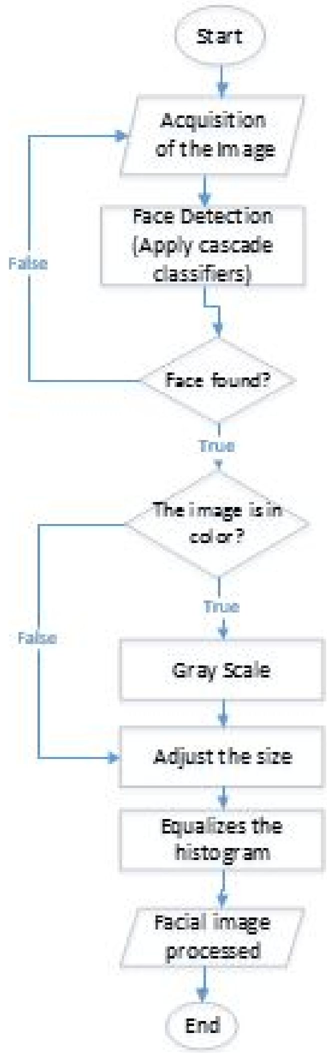


Figure 4.5: Flowchart reading and processing facial image.

4.1.2 Recognition

To perform the recognition of a subject in front of the camera of the MASHI robot, it is required to perform the training and processing of the image as explained above, then open the application "reconocimiento.py" (Figure 4.6).

This algorithm looks the most similar facial image within the training data, and identifies the person with the name with which their facial images were stored in the database. A window called "MASHI-FACE RECOGNITION" is opened and shows the results (see Figure 4.7)

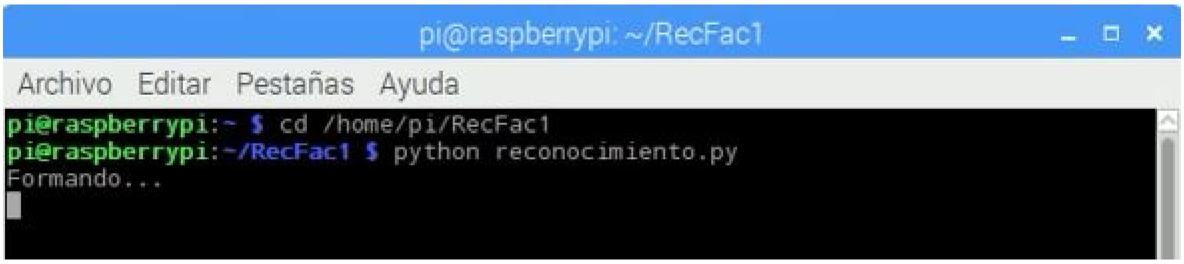


Figure 4.6: reconocimiento.py application.

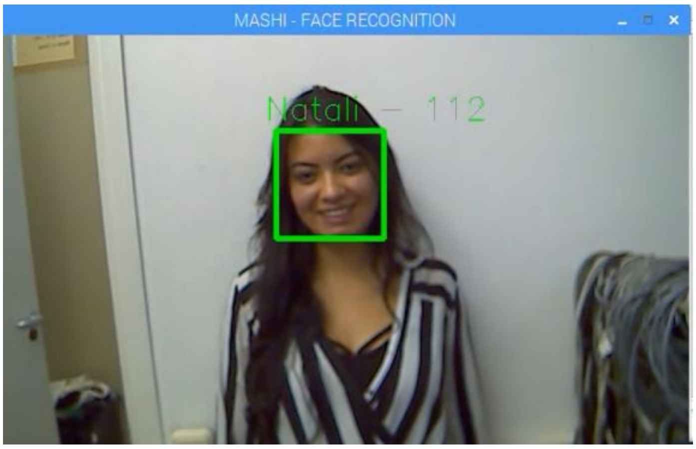


Figure 4.7: Face recognition with the application.

4.2 Experiments

This section describes the experiments performed to compare the efficiency of the face recognition algorithm implemented in the MASHI robot. Three types of experiments were carried out to evaluate the performance of the implemented solution. It is necessary to define the test scenarios that must be considered for the application of the MASHI robot. In that way, the proposed tests must be performed by varying some parameters that are important in the learning and recognition process of the solution.

1. Identification percentage.
2. Change the distance between the robot and the test subject.
3. Variations in illumination.

To develop the tests, 8 students from the Universitat Politècnica de Catalunya (UPC) were taken as a database. The test was performed and the efficiency percentages were analyzed in each case. Figure 4.8 shows the subjects with their respective IDs. After performing the training of the MASHI robot with the test subjects, facial recognition is performed. The possible results provided by the system are described below:

Figure 4.8: Test subjects.

1. **True positive:** This occurs when the person's information is in the database and the person identified corresponds to the one that is in the training database. Figure 4.9 shows the result after applying facial recognition, in this case, the result is correct.

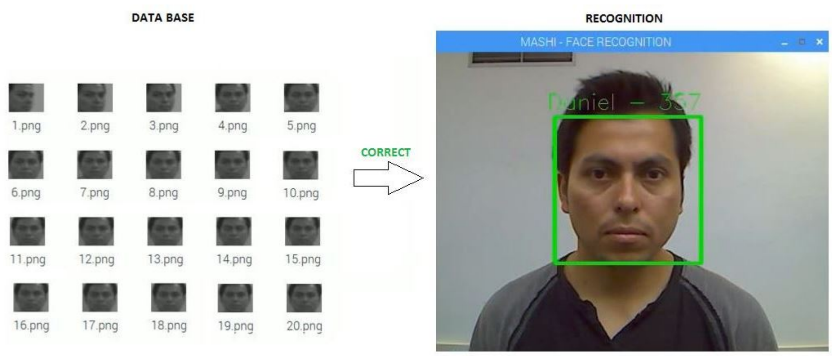


Figure 4.9: True positive.

2. **False negative:** This happens when the person's information is in the database, and the system cannot identify the person. Figure 4.10 shows the result after applying facial recognition, in this case, the identification label is "unknown", and thus the result is incorrect.

3. **False positive:** This occurs when the person's information is not in the database, and the system identifies the person with another face that is already included in the database. Figure 4.11 shows the result after applying facial recognition, and in this case, the result is incorrect.

4. **True negative:** This happens when the person's information is not in the database, and the system cannot identify it. Figure 4.12 shows the result after applying facial

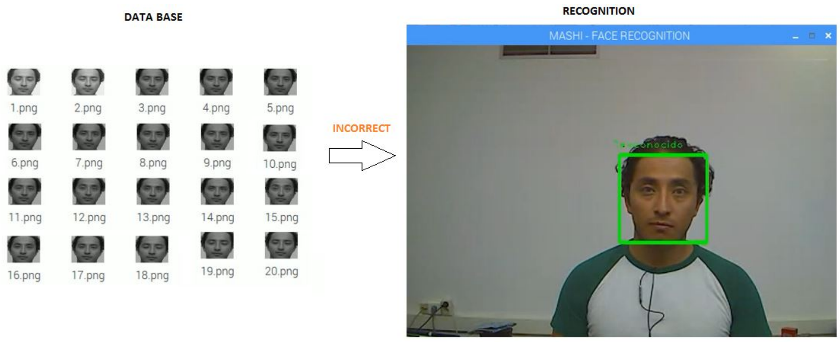


Figure 4.10: False negative.

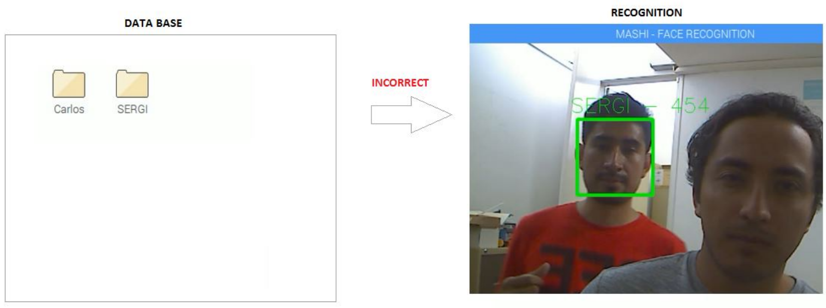


Figure 4.11: False positive.

recognition, in this case, the identification is “unknown”, and thus the result is correct.

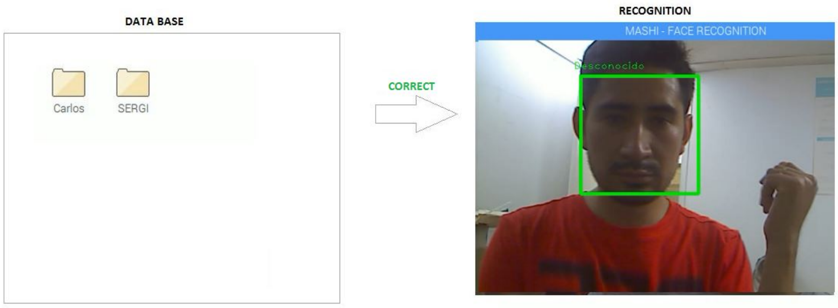


Figure 4.12: True negative.

4.2.1 Identification percentage

4.2.1.1 Scenario

To define the first scenario (see Figure 4.13):

- The training will be done by leaving a separation distance of 100 cm between the robot and the human. Moreover, twenty images will be taken per person.
- The recognition will be performed 5 times per person, at the same distance as the one used during the learning process. However, the threshold will vary to identify the best value of the threshold to carry out the best rate of face recognition.

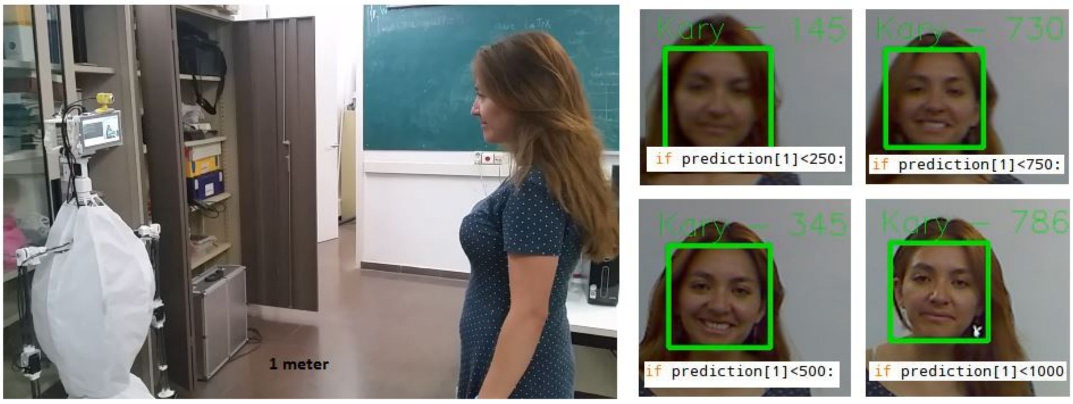


Figure 4.13: Identification percentage scenario.

The following Table 4.1. shows the set parameters to develop the first experiment.

Parameter	Value
TRAINING	
Number of people [u]	4
Robot-Human distance [cm]	100
RECOGNITION	
Threshold	<250; <500; <750; <1000
Times	80 [5 each person]

Table 4.1: Parameters Experiment 1.

4.2.1.2 Tables, graphs and results

The Figure 4.14 shows a table of results, for each threshold value the row presents the number of subjects identified as true positive (TP), true negative (TN), false negative (FN) and false positive (FP)

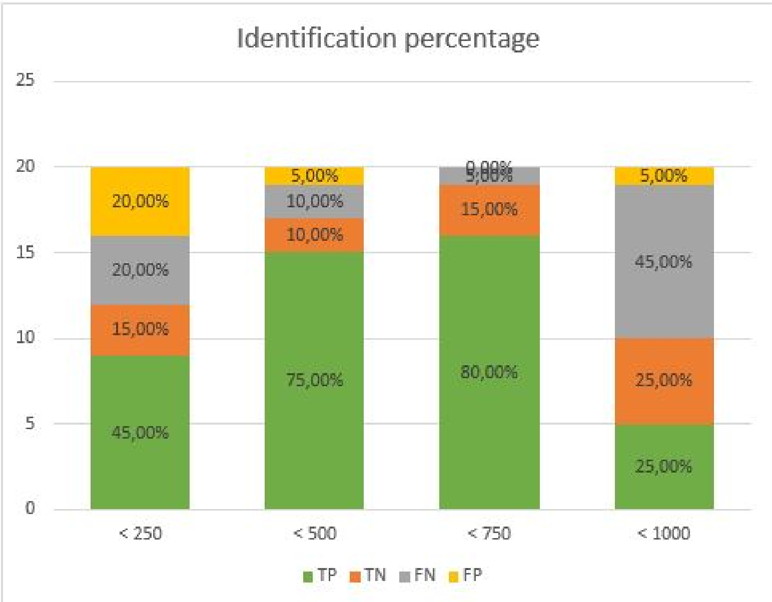
THRESHOLD	TP	TN	FN	FP	TOTAL
< 250	9	3	4	4	20
< 500	15	2	2	1	20
< 750	16	3	1	0	20
< 1000	5	5	9	1	20

Figure 4.14: Identification percentage scenario.

Figure 4.15 shows the percentage of recognition considering 20 attempts to identify as 100% for each case

THRESHOLD	TP	TN	FN	FP
< 250	45,00%	15,00%	20,00%	20,00%
< 500	75,00%	10,00%	10,00%	5,00%
< 750	80,00%	15,00%	5,00%	0,00%
< 1000	25,00%	25,00%	45,00%	5,00%

(a) Percentage table.



(b) Percentage graph.

Figure 4.15: Identification percentage.

4.2.1.3 Analysis

By observing the resulting graphs, it can be concluded that normally a true positive is when the value of the threshold is less than 750 and a false negative is when this value is less than 1000. Between 500 and 700 the algorithm sometimes recognizes the person accurately, but other times it confuses some faces.

4.2.2 Changing distance between the robot and test subjects

4.2.2.1 Scenario

To define the second scenario, the best result of threshold obtained in the first scenario was used, followed by the process of learning and recognition with different people.

- The training will be performed at distinct distances. Each white mark on the floor represents an accrual distance of 50 cm. (see Figure 4.16)

- The recognition stage will be performed with a variable distance to identify the working range of the system. For that process, it will be necessary to work with four people to obtain the different operation results (TP, TN, FN, FP). (see Figure 4.17)

Figure 4.16: Distances scenario.

Table 4.2. shows the set parameters to develop the second experiment. Furthermore,

Parameter	**Value**
TRAINING	
Number of people [u]	4
Robot-Human distance [cm]	50,100,150,200
RECOGNITION	
Threshold	<600
Times	36 [9 each person]

Table 4.2: Parameters Experiment 2.

the results obtained on this test must be compared with results of "Proxemic Behaviour" shown in [14]. In that way, the optimal range of operation regarding facial recognition for the MASHI robot can be proposed

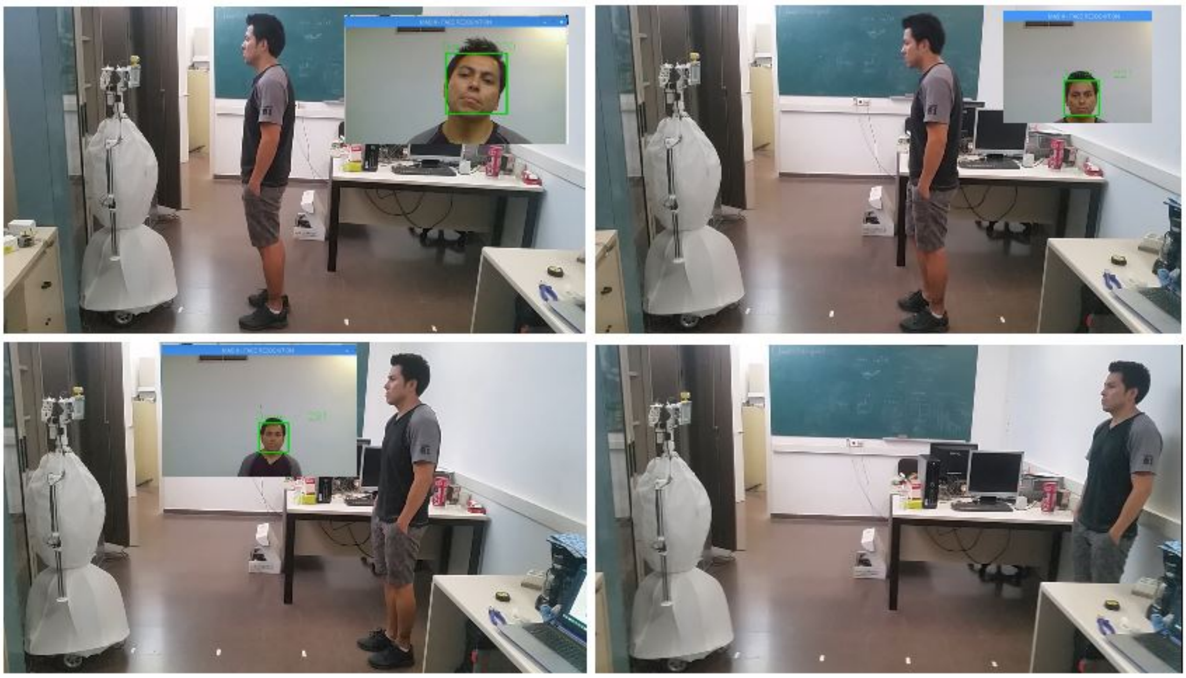


Figure 4.17: Recognition by changing distances.

4.2.2.2 Tables, graphs and results

Figure 4.18 shows a table with the results of the recognition at different distances. Figure

LEARNING	Distance [cm]	RECOGNITION 50	100	150
TP	50	4	2	0
	100	2	4	3
	150	1	3	4
FN	50	0	2	2
	100	0	0	1
	150	3	1	0
TN	50	0	0	0
	100	2	0	0
	150	0	0	0
FP	50	0	0	2
	100	0	0	0
	150	0	0	0

Figure 4.18: Results considering different distances.

4.19 shows the total results and the total percentage of TP, FN, TN and FP of this experiment.

4.2.2.3 Analysis

The performance of the algorithm when the training and recognition are performed by varying the distance between 50 cm to 150 cm are shown in Figure 4.19. When training

Distance [cm]	TP	FN	TN	FP
50	6	4	0	2
100	9	1	2	0
150	8	4	0	0

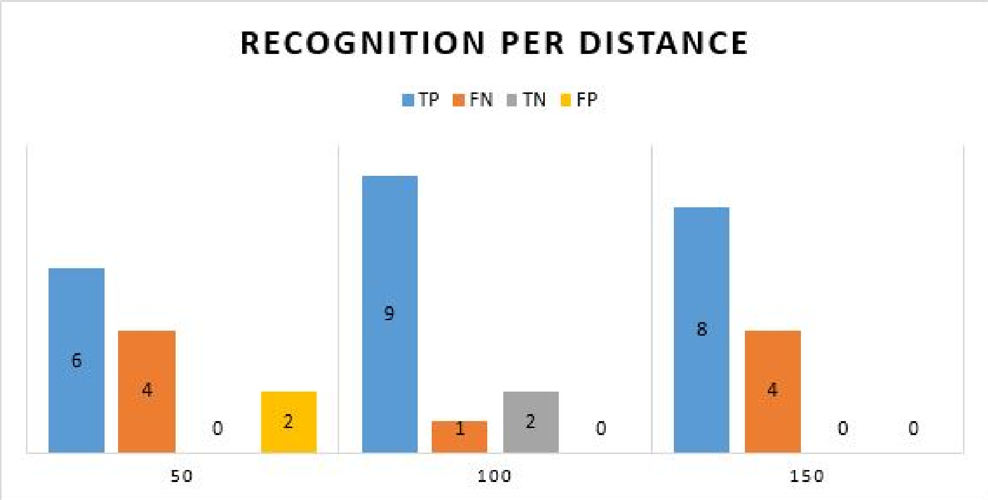


Figure 4.19: Results experiment two.

was done at 100 cm, the rate of recognition was bigger than for other training distances. Moreover, the highest number of recognition was precisely when the distance of recognition was 100 cm. However, it is important to notice that 73% of the recognition was obtained in the range between 100 to 150 cm. If we contrast the current results with the information of the study [14], regarding with the proxemics behavior of MASHI, the distance is adequate to keep a personal and social interaction with the humans. Where the yield of the platform reaches a value near to the 70% of reliability in terms of recognition. However, it is important to mention that the recognition does not work properly for distances over 1.5 m from the robot.

4.2.3 Variations in illumination

4.2.3.1 Scenario

In this scenario, the threshold and range of operation will be fixed according to the results obtained in the two previous experiments.

- For the training operation, the distance between the robot and the human will be 100 cm, the threshold smaller than 600, and the variable will be three different types of environmental lighting: natural light, artificial lighting, and low light (see Figure 4.20)
- In the recognition stage, the optimal distance will be used even if the light level that

is received by the camera lens varies. The process will be performed with four people to obtain the value of the illumination at which the algorithm works efficiently.
To measure the light around the camera, a Samsung application called "lux meter" was used, and two measures were considered: the vertical and horizontal component (see Figure 4.21)

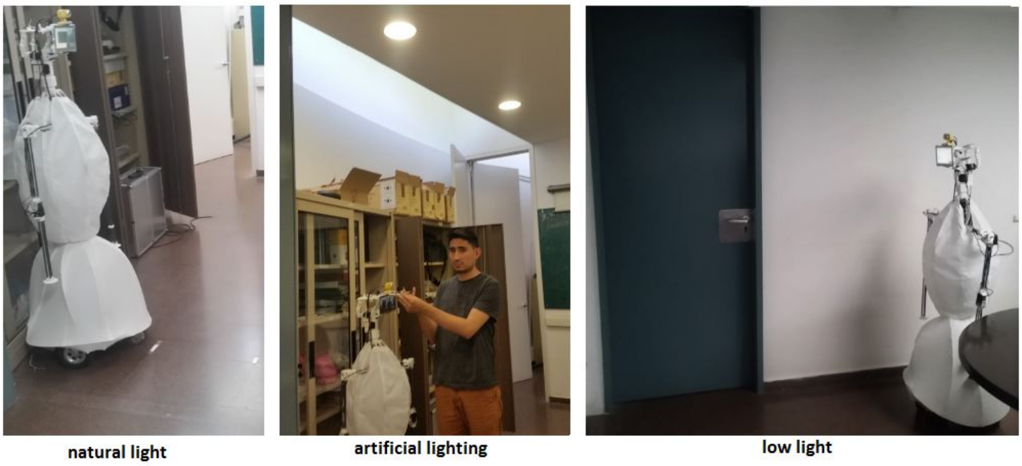


Figure 4.20: Illumination's scenarios.

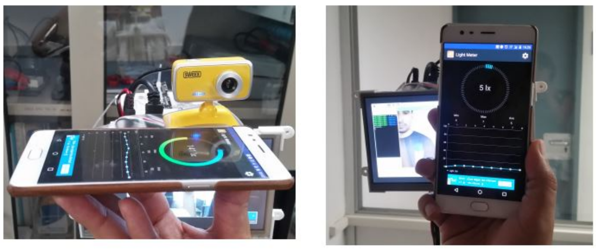


Figure 4.21: light meter app.

The following Table 4.3. shows the set of parameters chosen to develop the third experiment.

4.2.3.2 Tables, graphs and results

The results for the three previous environments are summarized in Figure 4.22, and Figure 4.23 shows the same information in percentage for each scenario.

4.2.3.3 Analysis

According to the evaluation standards of face recognition systems, the efficiency in a system is obtained when it has a maximum number of TP and TN and when the number

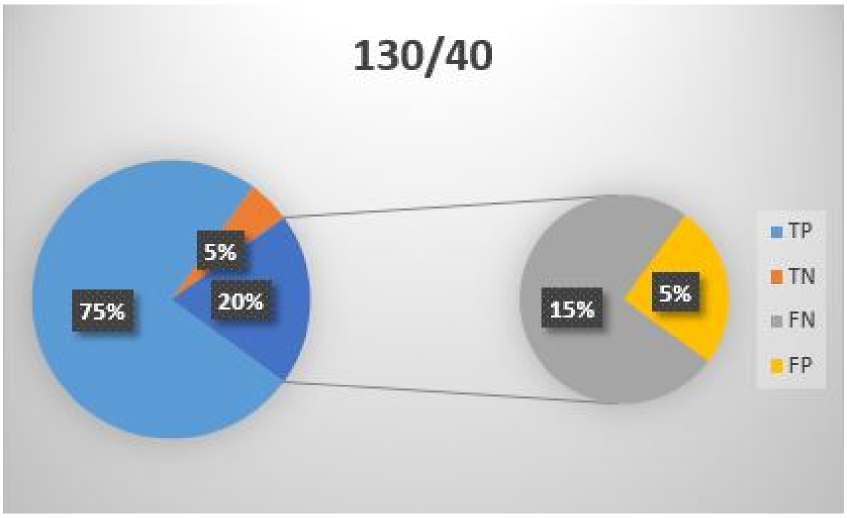
Parameter	Value
TRAINING	
Number of people [u]	4
Robot-Human distance [cm]	100
RECOGNITION	
Illumination	V:130/H:40; V:700/H:150; V:9/H:5
Threshold	<600
Times	60 [5 each person]

Table 4.3: Parameters Experiment 3.

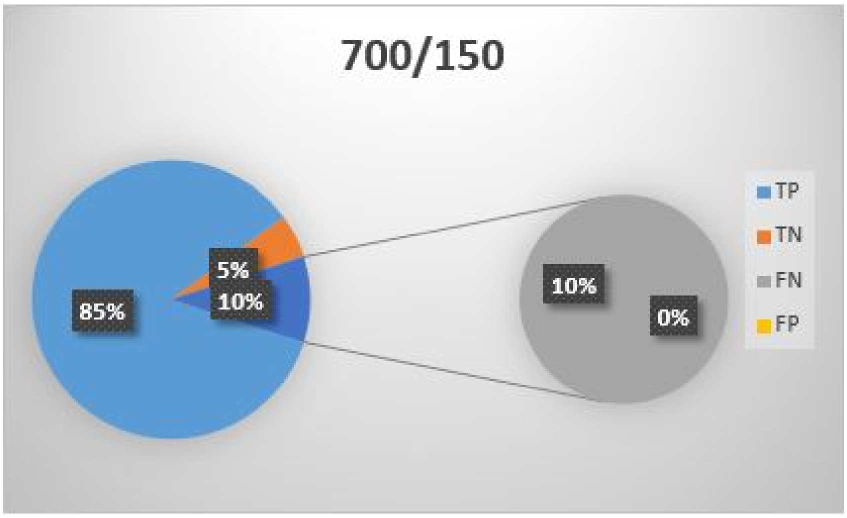
	LUX		RECOGNITION			
	V [lx]	H [lx]	TP	TN	FN	FP
LEARNING	130	40	15	1	3	1
	700	150	17	1	2	0
	9	5	11	4	8	1

Figure 4.22: Identification illumination's scenarios.

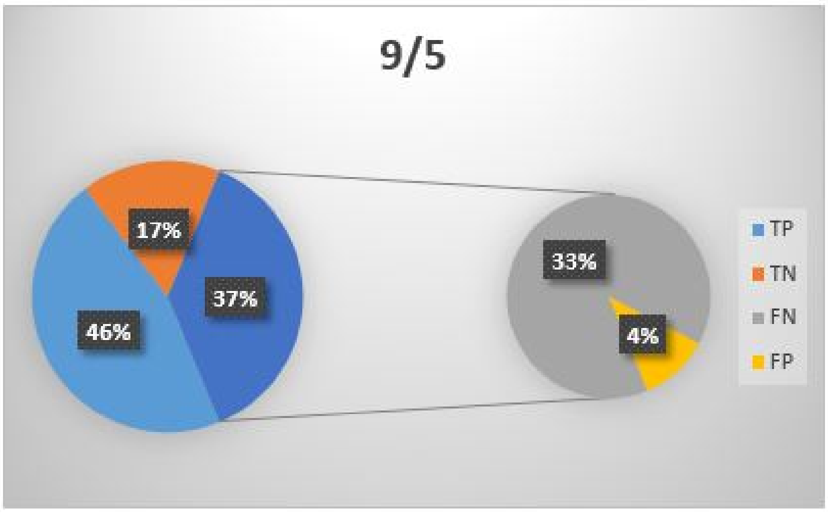
of FP and FN tends to zero. Therefore, by observing the results on the graphs it can be concluded that the system works the best with artificial light (90%) and with natural light (80%). The percentage of identification is low when MASHI works in a dark environment (63%).



(a) Natural light.



(b) Artificial lighting.



(c) Low light.

Figure 4.23: Results experiment three.

Chapter 5

Costs

5.1 Budget

The total cost corresponds to the Project "Facial Recognition System applied to Multipurpose Assistance robot for Social Human-robot Interaction (MASHI)" has been estimated at the equivalent of €13.489 (Thirteen thousand four hundred and eighty nine euros), based on the calculation of current unit prices in Spain on February 2017.

5.1.1 Structure cost

The prices of the different items of the material to form the structure of the MASHI are shown below in Table 5.1.

Material	**Cost**
Wood base	20
Vertebral column	20
Plastic rods and gears	30
Metal rods	20
Shaft collars	13
PLA	75
Aluminium profile	2
Wiring/Screws/Foam	40
Accesories	20
TOTAL	**240**

Table 5.1: Detailed structure cost.

5.1.2 Material cost

The price takes into account the unit costs of the material at distributor's commercial price; the following Table 5.2. shows the detailed pricing.

Concept	Part Number	Units	Price	Cost
Motor Mount and Wheel Kit	Parallax 27971	1	300	300
Motor Controller	Parallax HB-25	2	50	100
Motor Encoder	Parallax 29321	1	40	40
External Battery 12VDC	EnergiVm NHV1290FZ	1	30	30
Wheel's Micro controller	Arduino Mega 2560	1	40	40
Protoboard	Caolator MB-102	1	5	5
Servos' Microcontroller	Robotis OpenCM 9.04A	1	20	20
Display	AT070TN90	1	75	75
Front Web-Cam		1	30	30
Selfie Web-Cam	Logitech Inc WebCam C270	1	35	35
Speakers	Energy Sistem Music Z30	1	35	35
Servomotors	Dynamixel AX-12A	6	68	408
Motherboard	Raspberry Pi 3 Model B+	1	49	49
Transparent Case + heat sink	Aukru	1	7	7
Motor Controller	Parallax HB-25	2	50	100
Portable Charger 5V	iMuto 5000mah 2.1A	1	15	15
Memory card, 90 MB/s,U3	SanDisk Extreme 32 GB	1	20	20
TOTAL				**1,209**

Table 5.2: Detailed material cost.

5.1.3 Personnel cost

For the case of personnel costs, the current salary scale for public sector employees in Spain for 2017 is used as base; it is considered that an engineer earns 20 euros/hour and also that the time to develop the project was 25 weeks, 4 hours each day. Furthermore, the use of the equipment like a laptop and a 3D printer is considered too. The total personal expenses are summarized in Table 5.3:

Concept	Hours	Hour price	Cost
Engineer	500	20	10000
Laptop	500	0.2	100
3D Printer	60	3	180
TOTAL			**10280**

Table 5.3: Detailed personal cost.

5.1.4 Overall cost

Likewise, provision has been made for contingencies, estimated at 15% of the costs described above. The following Table 5.4 shows the overall cost of the project.

Description		Cost
Structure Cost		240
Material Cost		1209
Personal Cost		10280
	SubTotal	**11729**
Provisions (15%)		1760
	TOTAL	**13489**

Table 5.4: Overall cost.

5.2 Time

This project has been completed over a period of 25 weeks. During this time the phases of maintenance, reconstruction, programming and testing over the MASHI using a Raspberry Pi 3 have been developed, since the robot had to have the same functions as when it worked with a PC.

5.2.1 Maintenance and reconstruction tasks

- Some pieces were tested, these pieces allow the movement of the arms and head, and some of it was printed again and reassembled. Hence this work helped to find some errors in the initial design and to suggest some solutions.
- The encoders were changed in the base of the robot since the ones that were installed no longer worked.
- The change of the wiring of the structure was made, the cables were correctly labeled, and the designs were updated digitally with all the changes.

5.2.2 Programming tasks:

- The JAVA code was revised, and the study was made for the operation of the interface that enables the teleoperation over MASHI, and the application that allows upload photos on Twitter and interact with social networks.
- To perform the facial recognition, one of the main tasks was the installation of the OpenCV library over the Raspberry Pi 3. This task required a considerable time of the project. After programming the code to learn faces and for the face recognition was achieved.

5.2.3 Testing and report:

- Tests with different environments of the application made.

- All work has been fully documented

The Gantt diagram of Figure 5.1 shows the distribution of tasks and the relationship between them.

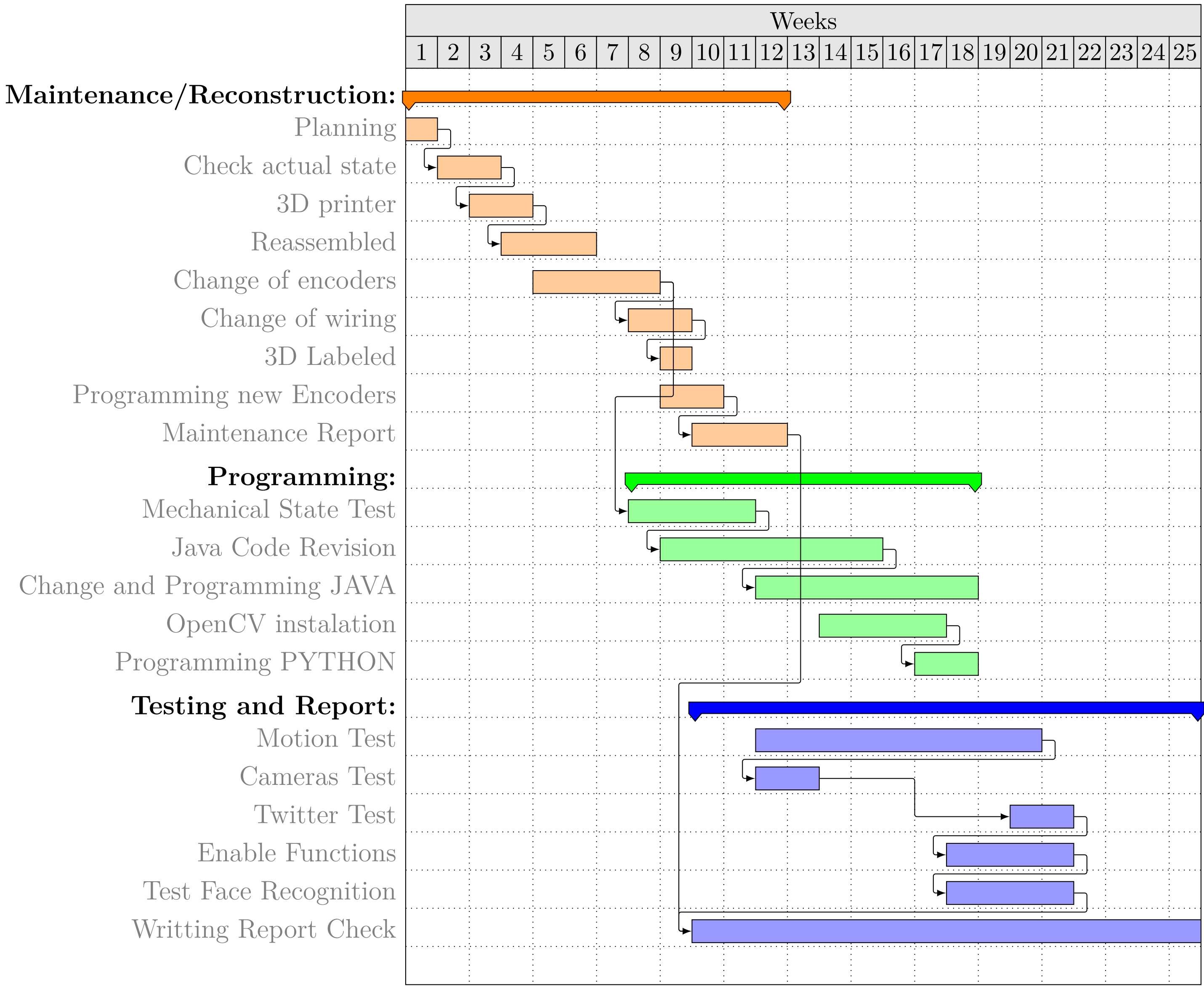


Figure 5.1: Gantt diagram.

Chapter 6

Environmental impact

The continuous increase in energy demand, the current global economic framework, and the strong push of the use of open source-based devices have led to the emergence of technological solutions that allow the optimization of energy consumption. These projects are very important to achieve both, cost reduction and environmental impact reduction, thanks to increasingly sophisticated and economic systems.
In the case of MASHI, the analysis of the environmental impact was done by calculating the CO_2 [10] produced by all the devices that make it up, and assuming that they are all working at the same time for one hour. Table 6.1 shows the power consumed.

Component	**Quantity**	**Power [W]**	**Total Energy [kWh]**
Servomotors Parallax	2	18.25	0.0365
Parallax HB-25	2	1	0.002
Dinamixel AX-12	6	10.8	0.0648
Arduino Mega 2560	1	0.25	2.5×10^{-4}
Robotis OPenCM 9.04A	1	0.2	2×10^{-4}
Display AT070TN90	1	7	0.007
Speaker	1	3.9	0.0039
Camera	2	1	0.002
Raspberri Pi3	1	4	0.004
TOTAL			**0.12065**

Table 6.1: Power consumption.
Data of power obyained from [10] X. Rodríguez Thesis

To calculate the "carbon footprint", i necessary use the methodologies recommended in the O.S.E manual according to ISO 14067 or the GhG Protocol [24]. The "carbon footprint" is calculated multiplying the kWh consumed by the emission factor of the electric energy FE = 0.39 Kg CO_2 / kWh [24]; the result is shown below:

$$0.12065\text{kWh} \times 0.39\frac{\text{Kg } CO_2}{\text{kWh}} = 0.047054 \text{ Kg } CO_2$$

Another impact to consider is the one caused by transporting the MASHI robot to the different presentations or expositions, considering the use of a car during one hour at an average velocity of 80 Km/h [24] The result is calculated with the following equation:

$$\text{Source of emission} \times \text{Emmision Factor} = \text{ Kg } CO_2$$

$$80\text{km} \times 0.197\frac{\text{Kg } CO_2}{\text{km}} = 15.76 \text{ Kg } CO_2$$

For the total calculation of the carbon dioxide emissions of MASHI over the course of one year, it is assumed that it will be used for at least four months for different practices lasting 3 hours per day, and also that it will be transported to two expositions per year. The result is shown in Table 6.2:

Description	**Time[h]**	Kg CO_2	**Total Emission** [Kg CO_2]
MASHI Operation	240	0.047054	11.29
Transportation	2	15.76	31.52
TOTAL			**42.81**

Table 6.2: Total emission of CO_2.

With the measurement of the levels of carbon dioxide emitted by MASHI in full operation it can be concluded that this robot has a low the air-quality impact by chemical load ,and that can be used in buildings with enclosed spaces and occupied areas [26].

Chapter 7

Conclusions

By accepting to the fact that a social robot like MASHI must be able to make interactive exchanges with its environment in a natural way, the communication skill of the MASHI platform was improved. During the hereby presented work, the OpenCV computer vision library was installed on the Raspberry Pi 3 and a recognition system that allows the robot to learn the faces of the people with whom it interacts was developed. This algorithm is also responsible for recognizing the faces of people approaching the robot. After the implementation of this feature, the experiments made demonstrated the effectiveness of the applied technique. Hence, this allows that MASHI can make the detection of faces inside a building in real time.

From the results obtained in the analysis of the experiments, it can be concluded that the appropriate threshold value should be between 500 and 700 , since within this range the recognition rate is between 80% and 90% of effectiveness. Furthermore, considering that the MASHI robot is a social robot and that people perform activities in front of MASHI, the second test shows that the recognition has an effectiveness of 70% when people move up to 1.5 meters from the base of MASHI. It should be noted that the developed application is oriented to an uncontrolled lighting situation, and thus the efficiency of the recognition system with variation of illumination is 74%.

As social robots become more common in our everyday environment, their ability to create sustained relationships is increasingly significant. In this work, the importance of the interaction between humans and robots was analyzed. It was concluded that for robots to get involved in social interactions without problems, algorithms and detection technology are needed to allow robots to obtain data in a precise and direct way, but often these devices reach very high costs due to the accuracy they require in their measurements.

Chapter 8

Limitations

Choosing a coding language used for developing the facial recognition system on the Raspberry Pi 3 was complicated. Firstly, Java was discarded since the headers of the OPENCV library were not found by Geany. Secondly, C ++ was discarded since the development time in this language would be slower than Python. Finally, it was decided to use Python 2.7 because it is a programming language that makes artificial vision system's development more agile.

Chapter 9

Future Works

As a continuation of this thesis work and as in any other research project, there are several lines of research that remain open and in which it is possible to continue working. During the development of this thesis have emerged some future lines that have been left open and are expected to attack in the future; Some of them, are more directly related to this thesis work and are the result of questions that have arisen during the realization of the same. Others, are more general lines as an option to future works for other researchers. The following list shows some future works that can be developed:

- Improve the MASHI's morphology in arms and head, if possible, the design and material of the structure should be changed to a more robust one that allows greater support and safe continuous movements at a higher speed.
- Today MASHI use a raspberry Pi 3 as a controller, where the face detection and recognition properties are working properly. However, the visual functionalities could improve if gesture recognition is included in the robot to give capabilities of gesture imitation that could be showed on the screen of the robot.
- It is proposed the change of cameras by some capable of capturing at a higher rate of FPS, in this way applications or projects with image processing that can be developed will have greater precision.
- The OpenCV library installed in raspberry can be used for some applications, for example, MASHI may be able to use vision for navigation control through the marks recognition, it is possible to make applications in which MASHI can interact with children and can teach them the colors and numbers.
- Increase proximity sensors, in this way MASHI has greater autonomy in its movements.

Bibliography

[1] C. Bartneck, T. Nomura, T. Kanda, T. Suzuki, K. Kato, *A cross-cultural study on attitudes towards robots* Uin Proceedings of the HCI International, 2015.

[2] M. Guadalupe, *Sistema de Visión para un Robot Social* Escuela Técnica Superior de Ingenieros Industriales, 2010.

[3] A. Clabiorne, C.Stephenson, C. Atkinson, K. Courtemanche, *Sentience: The Coming AI Revolution and the Implications for Marketing* Kindle Edition, 2015.

[4] J. Markowitz, *Robots that talk and Listen Technology and Social Impact.* Walter de Gruyter, Inc., Berlin/Munich, 2015.

[5] S.Strohkorb, Chien-Ming Huang, A. Ramachandran, B. Scassellati, *Establishing Sustained, Supportive Human-Robot Relationships:Building Blocks and Open Challenges.*, Department of Computer Science, Yale University, 2012.

[6] M. Beiter, B.Coltin, S. Liemhetcharat, *An introduction to Robotics with NAO.* Aldebaran Robotics, 2012

[7] Tomás González, *Artificial Vision in the Nao Humanoid Robot.* Universitat Rovira i Virgili, 2009.

[8] S. Lenser and M. Veloso , *Visual Sonar: Fast Obstacle Avoidance Using Monocular Vision.* International Conference on Intelligent Robots and Systems, 2006.

[9] M. Díaz, A. Andrés, J. Casacuberta and C. Angulo , *Robots sociales en la escuela.* Sevilla, 2011.

[10] X. Rodríguez, *Rediscovering the experimental robotic platform MASHI.* Universitat Politècnica de Catalunya, 2017.

[11] J. Cortés, *Redescubriendo la plataforma robótica experimental MASHI* Universitat Politècnica de Catalunya ESAII, 2017.

[12] N. Almeida, *Report of research, maintenance and application tasks on the MASHI robot of the UPC.* Universitat Politècnica de Catalunya ESAII, 2017.

[13] Yog M. Shukla , A. Tamba , S. Pandey , P. Sharma, *A Review and Scope of Humanoid Robotics.* RACE, 2014

[14] D. Paillacho, C. Angulo and M. Dìaz, *An exploratory study of Group-Robot Social Interactions in a Cultural Center.* Universitat Politècnica de Catalunya, 2015.

[15] Y. Zhang, K. Hornfeck, K. Lee, *Adaptive Face Recognition for Low-Cost, Embedded Human-Robot Interaction.* Case Western Reserve University, 2012.

[16] Hornfeck, K., Zhang Y. and Lee K., *Philos: A Sociable Robot for Human Robot Interactions and Wireless health Monitoring.* Symposium on Applied Computing (SAC 2012), Riva del Garda, Italy, Mar. 26-30, 2012.

[17] Open Source Initiative, *www.opensource.org/licenses/bsd-license.php* The BSD License, 2003.

[18] J. Rodríguez, *Reconocimiento Facial* Universidad Politècnica de Madrid, 2014.

[19] M. Gevara, J. Echeverry and W. Ardila , *Faces Detection in Digital Images Using Cascade Classifiers* Universidad Tecnológica de Pereira, 2008.

[20] D. Platero, *Reconocimiento de Imàgenes faciales orientado a controles de acceso y sistemas de seguridad* Universidad Distrital Francisco José de Caldas, 2015.

[21] P. N. Belhumeur, J. P. Hespanha, D. J. Kriegman, *Eigenfaces vs. fisherfaces: Recognition using class specific linear projection* ECCV (1), volume 1064 of Lecture Notes in Computer Science, Springer, 1996.

[22] M. Sierra, *Estudio comparativo de modelos de identificación facial basados en correlación* Escuela Técnica Superior de Ingeniería. Universidad de Sevilla, 2015

[23] OpenCV documentation, *Face Recognition with OpenCV* http://docs.opencv.org/2.4/modules/contrib/doc/facerec/faceretutorial.html fisherfaces-in-opencv.

[24] Observatorio de Sostenibilidad de España, *Manual de cálculo y reducción de huella de carbono.* OSE, 2012.

[25] Valsaín , *Cálulo de Emisiones* http://www.mapama.gob.es

[26] M. Berenguer and F.Bernal, *Carbon dioxide in evaluating indoor air quality* Instituto Nacional de Seguridad e Higiene en el Trabajo, 2000.