POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Executive Summary of the Thesis

Odile: an Expressive Robotic Agent for Emotional Exchange and Information Sharing

Laurea Magistrale in Computer Science and Engineering - Ingegneria Informatica

Author: Erica Panelli

Advisor: Prof. Andrea Bonarini

Co-advisor: Federico Espositi

Academic year: 2022-2023

---

1. Introduction

Recent studies in social robotics have led to an evolution in the quality of human-robot interaction and demonstrated how a robot can effectively engage in a conversation, entertain, and help the people around it. The research focuses on finding mechanisms and configurations that enable to animate an object that is itself lifeless. Movement is one of the main characteristics to be considered to build meaningful interaction with other people. Expression through gestures is used consistently in many social contexts.

This thesis was aimed at realizing a robot capable of representing different states of mind through the movement of its articulations, enabling a gestural interaction with a person on different levels.

To achieve this, we have considered the main concepts of several performing arts and cinema. We used them to define forms and mechanisms that could be sensible and usable in different configurations. The robot created was called Odile, as one of the characters of the Swan Lake ballet, inspired by the gestures typical of dance. We set ourselves several constraints during the execution of this project. To ensure that the focus was on movement and non-verbal communication, we chose a robot form that was not human-like. We also did not want to use any other type of feedback, visual or sound. The proposed project was carried out within the framework of the Physical Metaverse [3], which aims to create a real-world embodiment experience. For this reason, we did not make Odile autonomous but fully controlled. In doing so, we emphasized the deformability of the structure, which can change configuration and represent different emotional states.

2. Background and implementation context

During the development of this project, we focused on the study of the interaction between two individuals using theatre, dance, animated films, and exhibition art as reference points. We used the concepts learned to explore the range of minimum movements required for a robot to elicit a meaningful response from people.

In the context of theatrical performances, we specifically studied the contexts of mimicry, figure theatre, and Commedia dell'Arte. We have chosen to focus primarily on these three themes. These topics provide core principles for studying easily recognizable expressive gestures.

We also considered the theory behind a dancer's movement. In particular, the Laban method [2] provides an effective schematization between gesture and emotion.

Our approach to animated films was profitable in understanding what is required to make a fictional character real. We applied these principles in the construction and the coding of the structure. Primarily, we used the twelve *Fundamental Principles of Animation* presented by Disney animators [5]. A study made by taking these rules and applying them to a semi-humanoid robot [4] was our inspiration for the mechanics implemented in Odile's joints.

During the design phase of the robot, we were confronted with the need to make the structure as favorable and inviting to interaction as possible while still maintaining its abstract form. In this sense, studying the concept of the Uncanny Valley [1] was relevant for us to avoid mistakes in the design.

Starting from the already defined Physical Metaverse framework, we used the concepts learned as background for Odile's implementation. We took cues from the results obtained during several experiments carried out with five robots designed to be used in this context.

These trials brought to light issues that, in some cases, prevented the establishment of an interaction between the person and the robot. We were able to identify three guidelines that we followed to improve the so-far obtained results:

- Activate joint attention by focusing the movement on one body part.
- Make the robot capable of doing expressive movements by making them as smooth as possible.
- Design the robot's appearance so that it invites people to interact.

3. Preliminary design phase

By summarising the concepts from the theoretical and experimental background, we can define the implementation constraints that we have considered during this thesis:

- The robot can and must have any shape. The aim is not to replicate human sensor system.
- The robot should include components that enable it to perceive the actions of those around it or even react in the event of obstacles.
- The structure must be articulated in such a way that the actions of the controller can be mapped to the external environment.
- The robot will not have to have predetermined configurations, but the control of movements will have to be left to the user impersonating the avatar.

Before the realization of the final structure, we carried out a preliminary design period to ensure that all guidelines were met from start to finish.

We started by designing a structure that was as generic as possible and could be modified according to the emotions communicated by the controller. After several design attempts, we decided on a module containing an articulation that could develop upwards. This decision was made to bring the robot's end-effector to a height comfortable to the eyes of the person in front of it, facilitating its interaction.

In addition to considering the direction of the movement and, thus, the end-effector mentioned above, we observed the need to integrate a component that would allow the robot to capture the gaze of the person with whom it is interacting.

With this in mind, we designed another joint representing the robot's eye.

4. Odile's structure

We started to design and then build the final structure of the robot based on what we had learned from the prototypes [Figure 1]. We constructed a modular design by segmenting it into three distinct modules: wheeled base, core, and expressive structure.

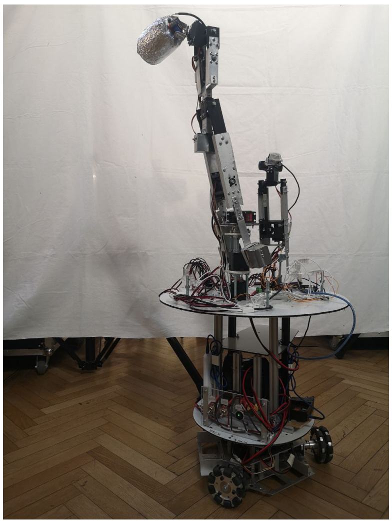

Figure 1: Odile's structure

The first and third modules are described in more detail below. The core level houses the robot's central controller, namely Raspberry Pi, designed to be connected to the lower and upper modules. It also comprises the batteries and dedicated switches for the separate power supply of the modules and the controller.

4.1. First module: Wheeled base

We used the larger base model already tried and tested in pre-existing robots. This base allows the weight of the entire robot to be stably managed while also providing ample play on the maximum speed attainable by the motors.

We reworked the chosen base to make it usable with the structure we had designed.

We used three DC motors positioned at a 120° angle to each other, forming a triangular base. We chose three omnidirectional wheels, which enable the robot to reach freely any pose.

4.2. Third module: Expressive structure

This module contains all the truly expressive and interactive parts of Odile. It includes two mechanical arms placed on a circular panel. One joint represents the end-effector necessary to drive the person's attention, and the other identifies the gaze. The structure has been designed to maintain the same recurring design in all joints.

The joint that provides the direction of movement is called Odile's tail. It is composed of six independent degrees of freedom designed to manage weight distribution:

- Base rotation: we designed a double-bearing mechanism that allows the arm to rotate smoothly by 180° without straining the servomotor at the base with the weight imposed by the upper joints.
- Vertical base tilt: we used counterweights to make it easier for the servo motor at the base to move the entire arm. We used a servo motor with a maximum torque of 60 kg/cm for a total arm length of 50 cm.
- Intermediate joint tilt: the dynamics of this joint are made possible through a belt drive; we used counterweights to make the dynamics smoother and avoided overloading the servomotor.
- Tilt-pan-tilt: this mechanism was taken from the design of the EMYS robot [4], which implements its construction in a pyramid shape. We adapted the servo motor configuration vertically to suit our case. This allowed a circular and expressive movement of the end-effector placed above the last tilt motor.
- End-effector: In the tail case, we completed the interaction by allowing it to receive tactile stimuli from outside. We connected the end-effector with a capacitive sensor (MPR121, Adafruit) to detect touch.

Next, we designed the structure of the second arm. This joint represents Odile's eye and gives her direction of gaze. We decided to keep a limb more elementary than the tail to differentiate the movement's direction from the eye.

Odile's head consists of four degrees of freedom:

- Vertical base tilt: This mechanism is designed in a mirror-image manner to the one used for the tail. The only difference lies in the weight management. Given the lighter structure, we decided to remove the counterweights and use a servo motor capable of supporting 25 kg/cm for a boom length of 30 cm.
- Tilt-pan-tilt: The structure realized replicates exactly that seen for the tail. We kept the same configuration to allow the head end-effector to make natural movements.
- End-effector: To make the role of the end-effector's eye of this arm explicit, we connected a webcam to the end. The webcam allows the robot controller to get visual

feedback on its surroundings.

5. Programming and testing

5.1. Movements

The programming of the modules under discussion here was done using the C++ language and the Arduino IDE software.

In this section, we illustrate the programming of the movements specifically for the third module. For each configuration, the aim was to find a representative behavior of a mood or an emotion and, at the same time, controllable from the outside.

The configurations implemented were:

- Independent control of each motor.
- Predefined movement sequences by varying the speed of the motors via a sigmoid function.
- Realization of a continuous biological movement representing a specific emotional state. In particular, we focused on two expressions: calm and agitation. To emphasize the tension, we have used a random step function. In contrast, we identified the mood related to calmness as generated by a sine function.
- Control of the end effector by inverse kinematics. This decreased the robot's degrees of freedom and simplified its control, but did not have the desired results due to the simplicity and consequent inability of the joints to reach all possible coordinates.
- Three different gestures to replicate the movement of the spinal column. We identified three configurations (forward-backward, up-down, and left-right) that could make the two arms easier to control by limiting their degrees of freedom and being expressive.

5.2. Control

The different controls implemented were advantageous for both prototyping and performing experiments. We tested the various control modes, looking for the simplest and most effective one.

Bringing the discussion back to the Physical Metaverse project, we used a series of implemented controls for remote communication. We found that the most precise and easiest method to use a controller is provided by the VR visor associated to its joysticks.

6. Experiments

Focusing on the main objective of this project, we defined two levels of interaction that can be experienced with the robot we built:

- Understanding. The robot must be able to communicate with a person by transmitting directions and collaborating.
- Empathy. The robot must be able to convey emotions and desires.

Given these premises, we wanted to split the test cases into two scenarios, keeping the same setting and only modifying the robot's behavior and goals.

We created a game to be played in pairs, the robot, and the test subject. The experience takes place in a maze with seven scattered buttons, some of which are hidden. The player should get out of the room as soon as possible, but can only do that after pressing three right buttons. Each time the subject presses a wrong button, the time will decrease significantly, so we wanted to force players not to choose randomly but to participate actively in the experience.

Before participating, people are warned that only the robot knows the position of the correct buttons. The realized maze was isolated to prevent other waiting participants from understanding the game solutions.

For each experiment, we found it appropriate to divide the test subjects into two groups: Alpha and Beta. The Alpha group allowed us to verify the robot's capabilities in the specific situation and and to define changes to the set-up to make the experiments with the Beta group profitable.

6.1. Environment set-up

The game environment consists of several elements: the room, buttons, a game manager, and the robot itself.

The maze [Figure 2a and 2b] is designed to have three levels of difficulty:

- The first is located at the entrance, where the robot has to prevent the player from pressing the wrong button and address him to the correct one.
- The second is where the robot has to point the test subject to the correct button out of three different ones.
- The third is that the robot must instruct

the test subject not to take a straightforward corridor to reach a wrong button, but to go around it, take a second one and reach the correct button.

(a) Realized maze (frontal)

(b) Realized maze (rear)

Figure 2: Realization of the experiment's maze

We implemented a game controller that could provide a visible user interface during the game. Inside it, a countdown was placed to highlight the remaining time, and three empty spaces were gradually colored green each time one of the correct buttons was pressed.

We used the game manager module to record each player's performance. We collected these data for the processing of posterior results.

We configured the robot's gestures using the proposed three degrees of freedom model, which emulates the movements of the spine. To control Odile, we used the mapping of the degrees of freedom to the VR visor and the corresponding joysticks. This configuration allowed linear control of the robot and was easy to learn.

6.2. First experiment's scenario

In the first scenario, we tested the robot's ability to communicate a goal, convey urgency, and dissuade the player from making mistakes. We used Odile's tail as a pointer without considering physical interaction.

Participants belonging to the Alpha group are defined as preliminary testers. We noticed a general negative trend in the comprehensibility of the robot's behavior and in its ability to express itself during the experience. After familiarising ourselves with the controls and standardizing the maze set-up by making some modifications, the results of the experiments with the Beta testers were more satisfactory. We can conclude that the negative performance was not due to the robot's conformation but solely to the controller's lack of experience and the confusing set-up of the initial environment.

6.3. Second experiment's scenario

The second experience is no longer collaborative, but the user plays individually. In this scenario, the robot's purpose has changed: Odile has to attract the player's attention and make him seek a caress. In this case, the robot must be able to express the need to set up a pact that makes it to provide support to implement the exit strategy if not caressed. Once caressed, the robot will show the player the correct button positions to get out. Odile's tail was used as a pointer and as a tactile tool.

The tests conducted on the Alpha group belonging to this scenario were negative on the whole. At the end of the preliminary tests, we concluded that the slow movements of the robot and its position close to the person helped the subject to establish a physical interaction through touch. We put the knowledge gained in the tests into practice with the Beta group of participants. At the end of all experiments, we observed an improvement in the players' evaluation of Odile's expressive abilities.

7. Conclusions

Investigating the domain of social robotics is an ongoing process, attempting to discover means to give life to mechanical forms of varying levels of complexity.

We aimed to create a robot that could relate to people on two levels. The first level is strictly communicative, and the second is purely empathic.

This goal was achieved during the experimental phase, as demonstrated by the presented results and the participation of the people who interacted with Odile. We succeeded in creating a

robot that was modular and adaptable to different types of emotional states.
Most of the problems encountered were due to the limitations of the robot's control. The number of degrees of freedom on the structure makes independent control of individual joints difficult.
Future developments concern the possibility of making the movements of the wheels and the expression module more direct and responsive by implementing trajectory controls and managing the weight distribution on the mechanical arms differently.
Another aspect to be explored is Odile's ability to respond to a person's spontaneous action. We still wonder whether the acknowledgment of the interaction can be activated solely by a movement or whether a more explicit signal is needed, such as a light or sound.

References

[1] Rina Diane Caballar. What is the uncanny valley? Creepy robots and the strange phenomenon of the uncanny valley: definition, history, examples, and how to avoid it, 2019.

[2] Todd Espeland. The eight efforts: Laban movement, 2023.

[3] Federico Espositi and Andrea Bonarini. Towards a framework for embodying any-body through sensory translation and proprioceptive remapping: A pilot study. In *xCoAx*, 2023.

[4] Ana Paiva and Tiago Ribeiro. The illusion of robotic life, principles and practices of animation for robots. In *HRI'12*, 2012.

[5] Frank Thomas and Ollie Johnston. *The Illusion of Life: Disney Animation.* Disney Editions, 1995.