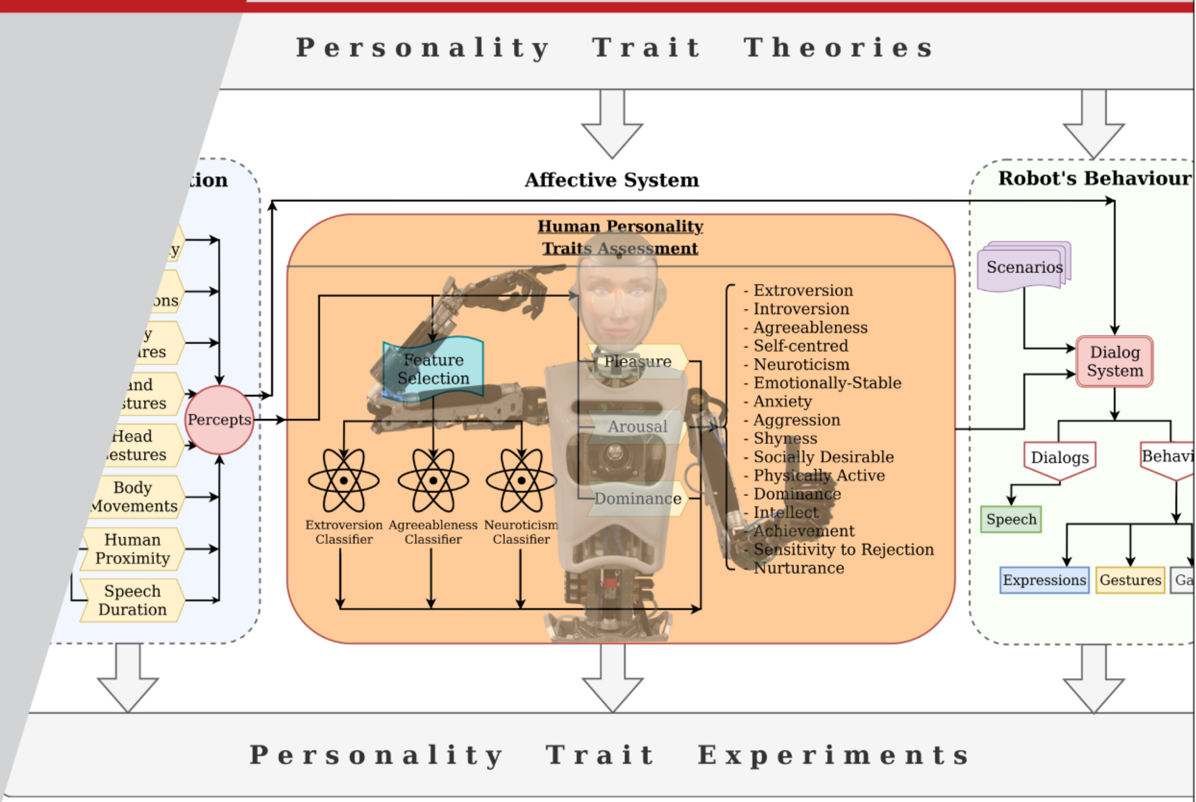
TECHNISCHE UNIVERSITÄT
KAISERSLAUTERN



Zuhair Zafar

Multimodal Fusion of Human Behavioural Traits: A Step Towards Emotionally Intelligent Human-Robot Interaction

RRLAB
THE ROBOTICS RESEARCH LAB

Multimodal Fusion of Human Behavioural Traits: A Step towards Emotionally Intelligent Human-Robot Interaction

Thesis approved by the

Department of Computer Science

Technische Universität Kaiserslautern

for the award of the doctoral degree

Doctor of Engineering (Dr.-Ing.)

by

Zuhair Zafar

Date of the defense	:	20.05.2020
Dean	:	Prof. Dr. Jens Schmitt
Chair of the committee	:	Prof. Dr. Reinhard Gotzhein
First reviewer	:	Prof. Dr. Karsten Berns
Second reviewer	:	Prof. Dr. Aleksandar Rodić

D 386

Acknowledgement

The success and final outcome of this thesis required a lot of guidance and assistance from many people, and I am extremely privileged to have got this all along with the completion of my thesis. All that I have done is only due to such supervision and assistance, and I would not forget to thank them.

I wish to express my sincere appreciation to my supervisor, Prof. Dr. Karsten Berns, who has the substance of a genius: he convincingly guided and encouraged me to be professional and do the right thing even when the road got tough. Without his persistent help and his undying interest in the field of human-robot interaction, the goals of this thesis would not have been realized.

I would like to thank Rita Broschart, for her administrative support throughout my stay in the Robotics Research Lab. She took care of all my contracts and administrative stuff without letting me worried a bit. Thank you, Frau Broschart for all the help.

I would like to acknowledge the efforts of my research partner and a good friend Salah Al-darraji in building a stable system. I have learned a lot from him during his stay in the Robotics Research Lab. My best wishes to him. I would also like to acknowledge my master thesis students, Daniel S., Sarwar, Rahul, and Ashita, for working with me. Their efforts have helped me in achieving the goals of this thesis.

I am grateful to all the members of the Robotics Research Lab for their support during my stay in the group. I would like to thank Massimo and Atabak for their technical as well as emotional support. I am grateful to Axel V. for helping me throughout the duration. I would also like to thank Hamza and Hannan for supporting me in difficult times. Tanittha, Qi, Alex, Sascha S., Patrick F., Steffen S., Thomas, Songlin and Sarwar, thank you, guys.

Special thanks to my parents for being patient with me during this long time. My father's support is the motivating factor in completing this thesis on time. I would like to recognize the efforts of my brother, Tayyab, for reading my thesis. His suggestions meant a lot to me. My special regards to my brother, Umair, and my sister, Rabbia, for their moral support during this time.

In the end, I would like to express my deep gratitude to my wife and my son. I am indebted to them for their support, especially during the last year. They have supported me through thick and thin. The thought of going home after a long tiring day refreshed me tenfold. Thank you for being there for me.

Kaiserslautern, February 2020
Zuhair Zafar

Abstract

With the technological advancement in the field of robotics, it is now quite practical to acknowledge the actuality of social robots being a part of human's daily life in the next decades. Concerning HRI, the basic expectations from a social robot are to perceive words, emotions, and behaviours, in order to draw several conclusions and adapt its behaviour to realize natural HRI. Henceforth, assessment of human personality traits is essential to bring a sense of appeal and acceptance towards the robot during interaction.

Knowledge of human personality is highly relevant as far as natural and efficient HRI is concerned. The idea is taken from human behaviourism, with humans behaving differently based on the personality trait of the communicating partners. This thesis contributes to the development of personality trait assessment system for intelligent human-robot interaction.

The personality trait assessment system is organized in three separate levels. The first level, known as perceptual level, is responsible for enabling the robot to perceive, recognize and understand human actions in the surrounding environment in order to make sense of the situation. Using psychological concepts and theories, several percepts have been extracted. A study has been conducted to validate the significance of these percepts towards personality traits.

The second level, known as affective level, helps the robot to connect the knowledge acquired in the first level to make higher order evaluations such as assessment of human personality traits. The affective system of the robot is responsible for analysing human personality traits. To the best of our knowledge, this thesis is the first work in the field of human-robot interaction that presents an automatic assessment of human personality traits in real-time using visual information. Using psychology and cognitive studies, many theories has been studied. Two theories have been been used to build the personality trait assessment system: *Big Five* personality traits assessment and *temperament framework* for personality traits assessment.

By using the information from the perceptual and affective level, the last level, known as behavioural level, enables the robot to synthesize an appropriate behaviour adapted to human personality traits. Multiple experiments have been conducted with different scenarios. It has been shown that the robot, ROBIN, assesses personality traits correctly during interaction and uses the similarity-attraction principle to behave with similar personality type. For example, if the person is found out to be extrovert, the robot also behaves like an extrovert. However, it also uses the complementary attraction theory to adapt its behaviour and complement the personality of the interaction partner. For example, if the person is found out to be self-centred, the robot behaves like an agreeable in order to flourish human-robot interaction.

Contents

1. Introduction

The idea of autonomous robots is as old as the work of Aristotle [Joachim 52], who already foresaw the chances of machines doing humans' work to save time and effort. The increasingly widespread use of advanced robots, more specifically social robots, is one of several phenomena expected to contribute to the technological post-humanisation of human societies. *"A social humanoid robot is an autonomous robot that interacts and communicates with humans by following social behaviours and societal norms expected by society"* [Feil-Seifer 05]. Indifference to industrial robots, a social robot is made for serving humans in their private life, which makes the task of interaction more challenging.

Humans use their emotional system to interact intelligently according to the environment. They assess each event based on different emotional categories, e.g., positive or negative, comfortable or uncomfortable, exciting or dull and change their behaviour according to the situation. This is a ground-breaking difference between human and robot. It goes without saying that typical robots lack intuition and perception. However, computer science, along with the technology, has persistently been trying to make robots more intelligent. To some extent, rigorous research activities have paved the way for robots to act and behave more intelligently. Theories from psychology and cognitive science, when implemented for the robot, have shown promising results in this regard.

With the advent of high-end computer systems and technology, the idea of social robots helping in our daily life is not limited to fictional films and novels. As artificial intelligence is growing at a tremendous rate, robotic systems also become much more intelligent every day. A lot of task-specific robots have already been presented which are used in assisting patients, visitors, elderly people, and so on. These autonomous robots can undoubtedly improve their ability to function in complex environments and to behave appropriately in partnership with people. Using the properties of natural intelligence as a guide, a robot's cognitive system would enable it to figure out what to do, whereas the emotion system would help it to do so more flexibly in complex and uncertain environments. Besides, it helps the robot to behave in a socially acceptable and effective manner with people [Uleman 08].

In order to adapt and respond to human behaviour in all its unpredictability, social robots must need to excel at reading human intentions, understanding their behaviours over time and predicting human actions for natural human-robot interaction. Furthermore, social robots need to express different natural behaviours and reactions appropriately during an interaction. They need to show empathy towards interaction partner to appear more human-like. Considerable research has been conducted to enhance the robot's social skills. However, only a few robots are able to interact with humans. Various challenges have been documented in the literature due to the complicated nature of human interactions. A social robot must have cognitive abilities similar to human to realise some kind of human-robot interaction. These cognitive abilities, which include recognition of nonverbal cues such as, human postures, gestures, expressions along with others, help the robot to perceive different human actions. In order to analyse human emotions and behaviour, a social robot must also have emotional abilities. This emotional capability enables a robot to understand human emotions and behaviours by using the fusion of different cognitive abilities; hence making the robot emotionally intelligent.

For social robots to have such capabilities, a complex perception system is required to undertake these tasks. Previously, there have been few humanoid robots presented; however, either they have limited perception abilities, or they have an application-specific perception system. Although some researchers have developed an efficient perception system that can recognise human actions, perceiving human emotions and behaviours in real-time is still a challenge and mostly limited only to psychological research. Recognising human emotions and behaviours is a critical aspect of natural human-robot interaction.

The motivation of this thesis is the contribution to this goal by combining state-of-the-art vision and machine learning methodologies with psychological and social studies. This thesis contributes to the extension of a perception system concept for a socially interactive robot by introducing emotional capabilities such as personality and temperament assessment to develop an intelligent social robot. This perception system focuses on non-verbal cues of communication in daily life due to the importance of these signals in overall communication.

1.1 Objectives

The primary goal of this thesis is to develop a system that enables a robot to synthesize an appropriate behaviour adapted to human profile (i.e., personality). Human personality is made up of the characteristic patterns of thoughts, feelings, and behaviours that make a person unique. Personality plays a vital role in human-human interaction as it guides the conversation towards a level of satisfaction and comfort for humans. Personality assessment is often termed as the backbone of high-level human perception system. Understanding and expressing appropriate emotions is a fundamental building block of this system. Developing a personality assessment system needs to consider three psychological and sociological aspects: multimodal communication, personality theories and robot's adaptability.

Psychologists have long studied human personality, and throughout the years, different theories have been proposed to categorise, explain and understand it. According to [Vinciarelli 14], the models that most effectively predict measurable aspects in the life of people are those based on traits. Trait theory [Costa 98] is an approach based on the definition and measurement of traits, i.e., habitual patterns of behaviours, thoughts and emotions relatively stable over time. Trait models are built upon human judgments about semantic similarity and relationships between adjectives that people use to describe themselves and the others. For instance, consider most of the people know the meaning of nervous, enthusiastic, and open-minded. Trait psychologists build on these familiar notions, giving precise definitions, devising quantitative measures, and documenting the impact of traits on people's lives [Costa 98].

There are many complex models available in psychology to assess the personality traits of humans. Myers-Briggs theory suggests that personality types can be recognised based on our preference to particular objects, ideas, facts etc. There are four pairs in this model, namely extroversion and introversion, sensing and intuition, thinking and feeling, judgment and perception [Myers 80]. Another theory which deals with the temperament proposes four basic personality types, namely sanguine, choleric, melancholic and phlegmatic. These categories are named after the bodily humours, which have a relation to our bodily fluids [Eysenck 85]. Bodily humours are described as liquids within the body and identified as blood, phlegm, black bile and yellow bile.

However, the most dominant theory in the research of personality traits has been presented by Costa and McCrae [Costa 76], known as the *Big Five* model. The model consists of five dimensions, namely extroversion, open to new experience, conscientiousness, agreeableness and neuroticism. As far as this model is concerned, each dimension is considered to be a continuum or spectrum, in which the extremes are quite distinct. In other words, a person is placed somewhere on the continuum of each dimension based on the individual scores. Verbal and nonverbal facets are taken into account to assess possible personality trait. Interestingly enough, a person can be placed in more than one dimension, but there is a dominant personality trait ingrained in each person [Jensen 16].

Assessment of human personality plays a significant role in achieving natural interaction between human and a social robot. The significance of personality in human-human interaction can be better exemplified by two renowned theories coming from the field of human psychology, i.e., *the chameleon effect* [Chartrand 99] and *the similarity-attraction theory* [Henderson 82]. The chameleon effect explains the non-conscious human tendency to passively mimic the behaviour of one's interaction partner in a social environment. In contrast, the similarity-attraction theory emphasises that humans are generally attracted to and prefer the company of others who maintain morals and attitudes similar to their own. For example, it is quite often observed that there exists a sense of shared personality among friends than among random pairs of strangers. Similarity-attraction theory can be observed in people with thought processes such as not feeling alone in their belief, or the ability to predict the future behaviour of similar people in order to access the "window of bias" for enhanced relationships and validation of attraction. Besides, people tend to change their behaviour according to their interlocutor behaviour. If he/she is talkative and expressive, one also tends to be more expressive. Therefore, assessment of human personality is highly crucial for a robotic system in order to interact and adapt naturally for intelligent human-robot interaction.

To the best of our knowledge, there exists no research that has explored assessment of human personality traits concerning the adaptivity and behaviour of a social robot in the field of human-robot interaction.

The major contribution of the thesis is the development of psychological theories and models of personality on a real social humanoid robot to assess standard personality traits as well as subtle personality traits in real-time. To build such a system, the following tasks need to be accomplished:

Visual Perceptual System

The visual perceptual system of the robot is responsible for perceiving, interpreting and making sense of the world. Analysis of human behavioural traits requires a variety of perceptive skills. These perceptive abilities, which include the understanding of human nonverbal cues such as emotions, gestures, postures, and many more over time, help the robot to perceive different human actions. Usually, humans detect and recognise different behavioural traits of their counterparts by analysing their nonverbal cues over time. Extrovert people tend to be quite active, and their excessive hand movements during conversations most often show confidence and control [Oberzaucher 08]. Similarly, body posture, hand gesture, facial expression, etc. play an important role in extracting the emotional state of an interlocutor. This makes the nonverbal cues the basic building block for the assessment of human emotional state. Several psychologists have pointed out the various combination of non-verbal cues that denotes specific human behaviour trait. Accurate and robust detection of nonverbal cues is highly vital due to their significance in human interactions. The work in this thesis is focused on perceiving human personality traits in a visual way. Therefore, perceptive abilities, such as hand gestures, body postures, human proxemics, body activity, head gestures and facial expressions are considered.

Personality traits assessment

In order to understand human actions or reactions, understanding of behavioural traits is an essential requirement. Personality is a significant part of behavioural traits that expresses the characteristics of individuals in different situations. Human personality is made up of the characteristic patterns of thoughts, feelings, and behaviours that make a person unique. According to renowned psychoanalyst Freud [Holt 89], personality is a result of childhood experiences that are consciously or unconsciously embedded onto a person during the developmental stages. The old but famous saying *"First impression is the last impression"*, although one of many clichés, is based on the fact that humans tend to evaluate personality and make an assumption of oneself as soon as they interact. Persons engaged in an interaction behave differently based on the personality types they possess and the overall environment in which they act.

Apart from general personality traits, subtle personality traits, for example, shyness, aggression, and so on, are not categorised in the mainstream personality theories by psychologists. According to Watson and Clark [Watson 85], extroversion can be subdivided into the more specific facets of assertiveness, gregariousness, cheerfulness and energy. Similarly, neuroticism can be subdivided to loneliness, anxiety and sensitivity to rejection while shyness is the part of introversion trait. These subtle traits are as significant as the personality types during an interaction. In this regard, Russell and Mehrabian [Russell 77] introduced pleasure, arousal and dominance as three independent emotional dimensions to describe people's state of feeling. Mehrabian has used this three-dimensional emotional space and has identified several regions that correlate to certain traits [Mehrabian 96].

Personality assessment system can provide a robot with one of the essential features that humans have, which can enhance the interaction. Assessing personality relies on psychological observations. A variety of cognitive skills contribute towards each personality type, and a person can have multiple personality traits at a given time.

Real world experiments

Finally, the developed personality trait assessment system needs to be tested and evaluated. To test a robot's personality trait assessment system, real-world interaction experiments between a human and a robot have to be achieved. The personality trait system of the robot is a complicated one, which is arranged in three levels. The first level, perceptual level, is responsible for enabling the robot to perceive, recognise and understand human behaviour in the surrounding environment. Several percepts are extracted in this level. The second level, affective level, is responsible for the assessment of personality traits using the percepts from the first level. The last level, behaviour level, helps the robot in using the information from the perceptual and affective level to behave in an intelligent manner such as by adapting behaviour to interlocutor mood. Different tasks are involved in each level; each has a specific goal. Evaluation of these tasks is the first step of the appraisal.

The overall system needs to be tested in an interaction scenario between a human and a robot. In this thesis, I have implemented and tested the personality trait assessment system on a social humanoid robot, ROBIN. ROBIN has an upper torso, two arms, and a head with an expressive face that can exhibit facial expressions. The primary purpose of this robot is to investigate social human-robot interaction.

1.2 Document Outline

Chapter 2 describes human-robot interaction field in the context of social robotics. It discusses the type of interactions between humans and robots. This chapter also presents different challenges associate with the field of human-robot interaction. Few state-of-the-art robots are also discussed with the focus on robot's adaptivity during interaction and what mechanisms has been used to achieve them.

Chapter 3 defines "personality" from psychology point of view. Different theories are presented in order to investigate the role of personality in human-human interactions. Since the thesis is focussed on the assessment of personality traits, this chapter talks about different theories, such as *big five* theory, that are categorised in a particular way for easier personality assessment. This chapter also discusses the temperament framework for subtle personality traits assessment from psychological point of view. P.A.D. emotional space has also been discussed in this chapter.

The conceptual design of the proposed personality traits assessment system for intelligent human-robot interaction is described in Chapter 4. The chapter highlights the problems and challenges with the current social robots for human-robot interaction. The chapter also discusses the significance of personality traits using psychology and cognitive studies. The design of the proposed personality trait assessment system is presented. The major components of the architecture are discussed. Furthermore, the chapter discusses the requirements and the tasks to develop personality trait assessment system using nonverbal cues.

Chapter 5 discusses the visual perceptual system of the robot. The proposed personality trait assessment system requires perceptual skills similar to humans. The perceptual skills of a social robot are the abilities that enable a robot to perceive, recognize, and understand human nonverbal cues in the surrounding environment in order to flourish natural human-robot interaction. These perceptual skills helps the robot affective system to recognize human personality traits. Different nonverbal cues, such as face detection, human localization, posture recognition, body movements detection, facial expressions recognition, hand and head gesture recognition, proximity and speech duration estimation are developed and presented in this chapter. For posture recognition, a novel method has been introduced that takes joint position of humans and convert them into meaningful angles for later classification. Individual experiments with their results for each of these tasks are also discussed. Finally, the fusion process of all detected percepts is described to construct a human descriptor.

The development of personality traits assessment system is described in Chapter 6. Since multiple personality theories have been implemented in this thesis, these theories are briefly introduced. State-of-the-art personality trait assessment approaches are discussed and their drawbacks are highlighted. The significance of nonverbal towards each personality trait category is evaluated and shown its importance using Pearson's correlation coefficient. Moreover, a detailed survey results are discussed that shows the importance of each nonverbal cue. Chapter 6 also describes the implementation of pleasure, arousal and dominance emotional space using nonverbal cues. The temperament framework is also presented that uses P.A.D. emotional space to estimate subtle personality traits of a person.

Chapter 7 discussed the results of several personality trait experiments done using the humanoid robot ROBIN. First of all, efficiency evaluation of different modules of the system are conducted to confirm that the system is real-time and can interact with humans effortlessly without any delays. Then, multiple experiments have been done to evaluate *big five* personality traits. The interaction scenarios are also discussed and the confusion matrix with the results is discussed. Another set of experiments are discussed to recognize subtle personality traits of human. The performance evaluation of temperament framework is also reported on a single subject level as well as with average recognition rates of all the subjects.

Chapter 8 summarises the results achieved in this thesis and discusses the realisation of the thesis objectives. An outlook on the future work in the field of human-robot interaction is presented.

Appendix A presents details of the humanoid robot ROBIN used in this thesis. Appendix B explains the robotic framework used to realize the objective of this thesis. Appendix C discusses the external libraries used in this thesis implementation for the assessment of human personality traits. Appendix D shows the dialogue system used to manage the dialogue between human and the robot and lists the dialogue scripts that are used in the experiments. Appendix E reports the questionnaire that is used to evaluate the personality trait assessment system by asking subjects their opinions, suggestions and drawbacks of the system.

2. Human-Robot Interaction

Human-robot interaction (HRI) is the science of studying people's behaviour and attitudes towards robots in relationship to the physical, technological and interactive features of the robots. The primary goal is to develop robots that facilitate the emergence of human-robot interactions that are at the same time efficient (according to original requirements of their envisaged area of use) but are also acceptable to people, and meet the social and emotional needs of their users as well as respecting human values [Dautenhahn 13]. Interaction, by definition, requires communication between robots and humans. Communication between a human and a robot may take several forms, but these forms are largely influenced by whether the human and the robot are close to each other or not. Thus, communication and, therefore, interaction can be separated into two general categories:

- Remote interaction — The human and the robot are not co-located and are separated spatially or even temporally (for example, the Mars Rovers are separated from the earth both in space and time).
- Proximate interaction — The humans and the robots are co-located (for example, service or social robots may be in the same room as humans).

Within these general categories, social interaction between robots and humans comes in the category of proximate interaction. Social interaction includes social, emotive, and cognitive aspects of interaction. In social interaction, humans and robots interact as peers or companions. Therefore, these robots are also termed as social robots. *A social robot is a physically embodied, autonomous agent that communicates and interacts with humans on an emotional level. It is important to distinguish social robots from inanimate computers, as well as from industrial or service robots that are not designed to elicit human feelings and mimic social cues. Social robots also follow social behaviour patterns, have various "states of mind," and adapt to what they learn through their interactions.*

As social robotics becomes more and more cross-disciplinary beyond engineering and computer science and draws more on the knowledge and resources from Human-Computer Interaction, the term Human-Robot Interaction slowly substitutes the term "social robotics."

HRI as a cross-disciplinary field lies between robotics, AI, cognitive science, (developmental) psychology, interaction design, biology and especially ethology. HRI investigates in the question of the human perception of robot systems, of user-friendliness, on the question of design (anthropomorphic, zoomorphic or functional robots) as well as ethical considerations [Kiesler 04] [Burke 04]. The primary purpose of HRI is to develop more natural and effective interaction between human and robot. HRI is a vast field that addresses the question of how humans and autonomous robots should collaborate and how robots should be useful for humans. Previously, robots were used in industries remotely without any significant human interaction, but nowadays they are found in most technically advanced societies whether in daily indoor activities or in military field for bomb detection or for helping humans with disabilities.

Challenges in the field of human-robot interaction field are discussed in the following section.

2.1 Challenges in HRI

There are numerous challenges associated with the field of HRI due to its diverse application scenarios. As mentioned by Tapus et al. [Tapus 07], several intriguing questions and problems must be answered in order to develop emotionally intelligent and robust social robotic systems. Some of these challenges have been discussed in following sections.

2.1.1 Multi-Modal Perception

Real-time perception of the environment and its understanding is an incredibly complex task for a robot's perception system. The idea of emulating human perception system is based on the fact that humans would accept a social robot in their society that interacts and communicates using the same perceptive channels as humans. However, emulating human perception system is extremely difficult with the existing technology. Humans can process multiple perception task in milliseconds and can react instantly.

To function in a complex and unpredictable physical and social environment, humans apply their physical and intellectual resources to realise multiple goals intelligently and flexibly. Two distinct and complementary information processing systems – cognition and emotion enable humans, achieving these goals by operating simultaneously. The cognitive system is responsible for interpreting and making sense of the world. In contrast, the emotion system is responsible for evaluating and judging events to assess their overall value with respect to the human (e.g. positive or negative, desirable or undesirable, hospitable or harmful, etc.). When operating in the proper balance, the emotion system modulates the operating parameters of the cognitive system and the body to improve the overall mental and physical performance of the human.

Because of the limitations in the existing sensors such as vision sensors and auditory sensors, the realisation of perfect human-robot interaction is an open challenge. Vision and speech processing in the unstructured environment are both significant challenges for real-time processing. Similarly, language understanding and dialogue systems between humans and robots remain complex research challenges. In addition, understanding the connection between visual and linguistic data and its fusion over time makes the whole perception extremely complicated considering that all the sensory processing needed in a low-latency time frame which is suitable for human-robot interaction.

Another critical aspect of human-robot interaction is the type of sensors used for perception tasks. In order to realise natural human-robot interaction, some researchers have conducted experiments in smart structured environments with multiple sensors mounted at different locations. However, the challenge always is to process all the sensory data and fused it in real-time. Also, these systems can not be used outside laboratory settings. Moreover, understanding the environment and objects in the surrounding along with humans is overall an open challenge.

Perception of human is more demanding than other objects in the environments. The reason for the complexity lies in human behaviours. Human behaviours are triggered based on several psychological motives. Motives involve the biological, emotional, social and cognitive forces that activate behaviour. Understanding these behaviours based on psychology findings is highly critical for a robot's perception system, and it is still an open challenge. This thesis uses several psychology theories and findings to perceive human personality traits in the context of human-robot interaction.

2.1.2 Robot Behaviour

Social robots are designed to interact with people in human-centric terms and to operate in human environments alongside people. Many social robots, humanoid or animal-like in form, engage people in an interpersonal manner, communicating and coordinating their behaviour with humans through verbal, nonverbal, or affective modalities. As can be seen in the following examples, social robots exploit many different modalities to communicate and express social-emotional behaviour. These include whole-body motion, proxemics (i.e., interpersonal distance), gestures, facial expressions, gaze behaviour, head orientation, linguistic or emotive vocalisation, touch-based communication, and an assortment of display technologies.

For social robots to close the communication loop and coordinate their behaviour with people, they must also be able to perceive, interpret, and respond appropriately to verbal and nonverbal cues from humans. They need to express different reactions or actions in the same way as humans. Expressing different behaviours similar to a human on a social robot depends highly on the robot's design and physical limitations. Apart from the robot's design, expressing correct behaviour is also extremely important for human-robot interaction. Robot developers need to use cognitive psychology theories to develop appropriate behaviours and actions for a social robot. Although addressed several times in the research, developing such a general framework that controls and activates appropriate robot's behaviour is still an open research challenge.

2.1.3 Robot Design

Robot design is an important aspect in the development of a social robot for human-robot interaction. The primary goal is to design these robots in such a way that they are accepted in the human society. There are several machines and AI interfaces presented in the literature which are used for interaction with humans. However, not all these robots have been focused on the physical design and abilities of the robotic system.

Since last decade, many robotic systems have been presented in the research community that are designed based on human appearance. The goal is to design robots that has similar degree of freedom as humans. Moreover, these robot are also able to express emotions and different nonverbal behaviours. Robot designers have been focussed on the appearance of the robot, especially android robots, in the recent times. A silicon material has been used to cover the face of the robot which acts as a natural skin. However, expressing subtle expressions accurately requires many degrees of freedom which is an open challenge in the field of robot designing.

2.1.4 Ethical Issues

It might seem, at first glance, that the design of social robots and other intelligent systems which have more human-like methods of interacting with users is generally to be welcomed. However, there are several important ethical problems involved in such developments which require careful consideration. Given present progress and promises in robotic systems that interact with humans in a complex and intimate fashion, there is a need to discuss and clarify as many as possible of these problems. More human-like interaction with robots may seem a worthy goal for many technical reasons, but it increases the number and scope of ethical problems. There seems to be little awareness of these potential hazards in current human-computer interaction research and development and still less in current research in Human-Robot Interaction.

Designers can and do force their view of what constitutes an appropriate interaction on to users. In the field of computer science, in general, there have been many mistakes in this area. Some researchers (e.g., Norman 1999) argue that there is a systematic problem. These mistakes might be destructive instead of beneficial for society. Therefore, there is a prominent role for ethicists to play in the design of Human-Robot Interaction.

2.2 State-of-the-Art HRI Systems

Numerous social robots have been presented in the research community with varying degree of abilities and characteristics. The main idea behind the development of such social robots is to achieve natural human-robot interaction. Most of the robotic systems that are presented in the literature are focused on the perception system, some are focused on human-like appearance, and few of them are focused on interaction and communication with humans. The overall goal of these social robots is to emulate human-interaction abilities and use similar channels to perceive and express behaviours. We review state-of-the-art HRI systems based on the degree of interactivity, physical appearance and perception systems.

2.2.1 Android Robots

As suggested from the name, the main focus of development in such robots is their human-like appearance. Roboticists are developing social robots that resemble humans in appearance in order to be accepted in human society and after that, can perform a wide variety of general tasks in human environments and interact with humans in a social setting. Examples include the Geminoids robots.

Geminoids are teleoperated android robots that appear similar in appearance to their source. Hiroshi Ishiguro lab has developed several android robots at ATR [Nishio 07]. Geminoid HI-2 is a copy of Hiroshi Ishiguro that has 50 degrees of freedom that allows him to behave naturally, as shown in Figure 2.1. Geminoid F is a female teleoperated robot with 12 degrees of freedom. Other geminoid robots include Geminoid DK, Kodomoroid, Otonaroid and Telenoid.

Figure 2.1: Android Robots: (f.l.t.r.) Geminoid HI-2, Geminoid-DK, Geminoid-F, Kodomoroid and Otonaroid.

In addition to geminoids, other androids such as Junko Chihira and Jiya Jiya has also been presented in the robotics community. Junko Chihira is an android robot developed by Toshiba, Figure 2.2. The robot has been used as a receptionist in a museum. Using Toshiba's speech synthesis technology, the robot can speak in three different languages and can also recognise speech in three languages. Similarly, Jia Jia is also a female android robot developed at the University of Science and Technology of China in Hefei, Figure 2.2. The robot has an exceptionally realistic appearance and can recognise speech. Although the perception system is quite limited, it can localise a person and the gender of a person.

Figure 2.2: Toshiba's female android robot, Junko Chihira, on the left and Chinese android robot Jia Jia on the right.

The main focus area of these android robots is the natural movements of the head, facial expressions and other body movements. Because of their complex design and many degrees-of-freedom, these robots generally have a simple perception system. Therefore, most of these robots are externally controlled and lack autonomous human-robot interaction. Few of them can interact with humans by recognising human speech which brings the aspect of autonomous human-robot interaction. However, the visual perception system is generally quite limited in these robots. Despite the complexities involved, some androids are developed for human-robot interaction. One such example includes the Erica robot.

2.2.2 ERICA

Since the last decade, many androids have been presented in the robotics community. Most of these robots are replicated on the appearance of celebrities and individuals. The concept of androids, working side-by-side with humans, is also fairly depicted in many films and television shows. Despite the excitement over the appearance of these robots, the major goal is to create fully autonomous interactive androids. ERICA is one of the latest interactive android presented in the research community.

"ERICA", an acronym for "ERATO Intelligent Conversational Android", named for the Japanese ERATO Ishiguro Symbiotic Human-Robot Interaction Project, a collaborative research effort between teams at Advanced Telecommunications Research Institute International (ATR), Osaka University, and Kyoto University, with the goal of developing a fully-autonomous android [Glas 16]. ERICA is an upper-body female humanoid robot with controllable face and torso as shown in Figure 2.3. Arms and legs are not actuated.

ERICA' body has 19 degree-of-freedom, as shown in Figure 2.4. Most joint controls are focused on facial actuation for natural expressions and speech. Eyes have 3 DOF's, and it can be controlled in yaw, pitch and convergence. Eyelids (upper and lower) and eyebrows (inner and outer) constitute 4 DOF's. The mouth also has 4 DOF's (mouth height, width, upper and lower corners of the mouth. Another 2 DOF's in the face actuate jaw and tongue. The skeletal body axes shown in black in Figure 2.4 are actuated. These axes provide 6 independent DOF's: waist yaw, waist pitch, synchronous vertical shoulder movement, neck yaw, neck pitch and neck roll. The actuators used in the robot are pneumatic actuators regulated by servo valves.

Figure 2.3: The android robot, ERICA [Glas 16]

A custom voice designed for Hoya's VoiceText software is used to perform speech synthesis. ERICA also has two cameras 1280 x 1024 pixel 30fps cameras mounted in its eyes. Another two omnidirectional condenser microphones are mounted in the ears. Despite cameras and microphones mounted onboard, ERICA currently uses external sensors on a wired network for human position tracking, sound source localisation and speech recognition due to the mechanical limitations. For the detection of human, ERICA uses ATRacker tracking system [1] with external vision sensors such as Microsoft Kinect 2, 2D laser range finders [Glas 09] and a network of ceiling-mounted 3D range sensors [Brščić 13].

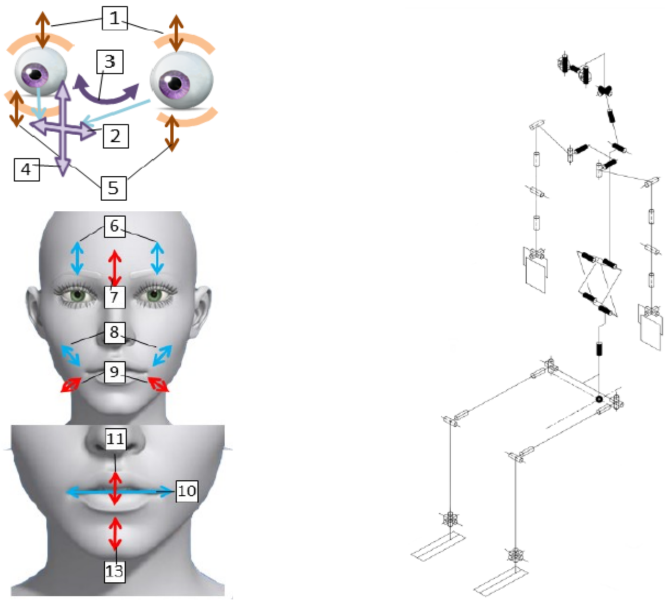


Figure 2.4: Degrees of freedom in Erica. Left: Facial degrees of freedom. Right: Skeletal degrees of freedom. Joints marked in black are active joints. [Glas 16]

ERICA uses two 16-channel microphone arrays, from which sound directions are estimated in 3D space (azimuth and elevation angles) with 1-degree angular resolution and 100ms time resolution. If detections by multiple arrays intersect in 3D space and a human is tracked at that position, a person is likely speaking. Afterwards, the input speech is enhanced by using the delay-and-sum beamforming. The automatic speech recognition (ASR) in ERICA is done by using the enhanced speech. To realise the speech recognition, the enhanced speech is processed by a de-noising autoencoder (DAE) to suppress reverberation components and signal distortion. The output speech signal of the DAE is decoded by an acoustic model based on a deep neural network (DNN). The DAE and DNN are trained by using multi-condition speech data so that it is robust against various types of the acoustic environment. Figure 2.5 shows the system architecture of ERICA robot [Inoue 16].

ERICA can also detect and identify whether a human is talking to her or any other person. ERICA tracks the user's location and head orientation in the 3D space by using the Kinect v2 sensor. The user localisation enables ERICA to spot if there is a person who wants to interact with her. ERICA identifies if the user is speaking to her by the head orientation. It enables ERICA not to respond when other people are talking with each other, for

[1] http://www.atr-p.com/products/HumanTracker.html

example, when a person introduces ERICA to a guest standing in front of ERICA. ERICA accepts user utterances when the following conditions are met: the user is standing in front of ERICA and looking at ERICA's face, and the sound source is coming from the direction of the user.

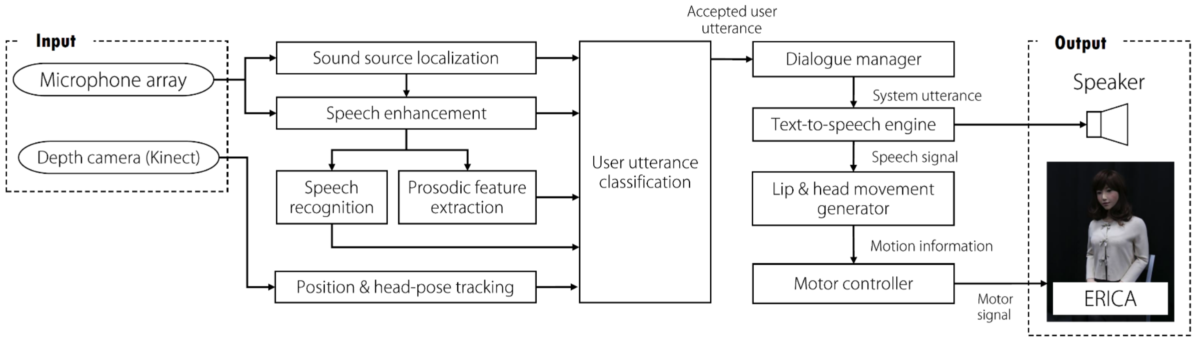


Figure 2.5: Multimodal Interactive System of ERICA [Inoue 16].

29 facial expressions and 16 gesture animations have been developed for ERICA. These included motions, such as bowing and nodding, as well as expressions of joy, sadness, surprise, disappointment, indecision, doubt, relief, wry humour, dissatisfaction, and many others. Together with these explicit expressions and gestures, the nonverbal behaviours have been used to generate breathing, blinking, gaze, backchannel nodding, and other "implicit" behaviours.

ERICA can speak and interact on various topics that are already designed offline. For evaluation, ERICA has been demonstrated publicly at several fairs. The primary interactive means used is verbal communication. The major shortcoming of ERICA is the use of *only* verbal channel for interaction. Vision-based perception of the scene and human understanding is quite naïve and limited only to human position and head tracking. Therefore, ERICA is unable to understand and react on nonverbal behaviours of users and can not adapt according to the mood and desire of the interlocutor.

2.2.3 ASIMO

ASIMO is a bipedal humanoid robot that has been developed by HONDA with a goal to develop robots that will coexist with and be useful to people since its first introduction in 2000. ASIMO, an acronym for "Advanced Step in Innovative Mobility", has been developed with the focus on two application directions: one is toward an "assistant robot" that helps people in their daily lives, and the other is toward robots that can substitute humans in areas that are dangerous or inaccessible [Shigemi 18]. Early development of ASIMO has been focused on solving complex challenges such as bipedal walking, running and jumping in different environments. The 2011 model of ASIMO can run at a speed of 9 km/h, run backwards, hop on one leg, hop on both legs and perform other such movements continuously.

ASIMO has overall 57 DOF's in its whole body. It has 3 DOF's in its head, 7 DOF's in one arm, 13 DOF's in the hands, 2 DOF's in the hip and 6 DOF's in one leg. ASIMO also has communication capabilities in order to interact with humans. It can recognise persons'

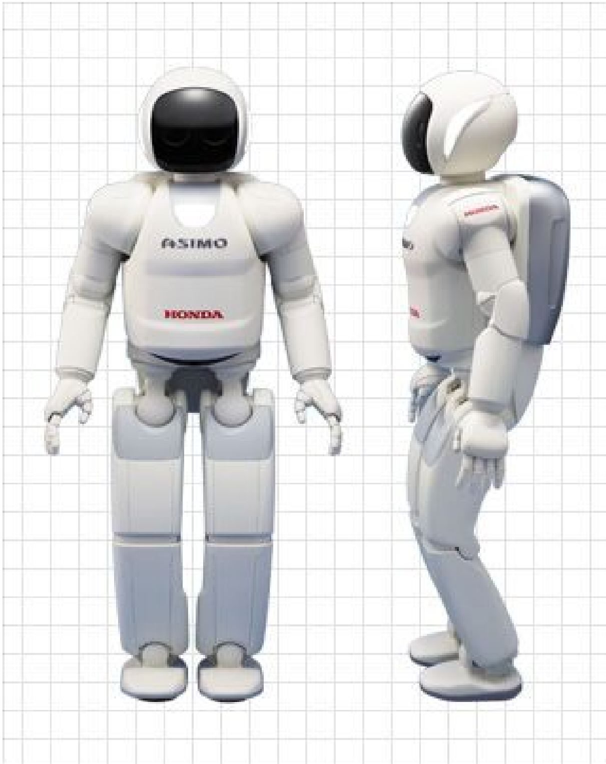


Figure 2.6: The humanoid robot, ASIMO, developed by Honda.

voices in 1-2 m distance. ASIMO uses 8-channel microphone array which is mounted around the robot's head. Open-source robot audio software [Nakadai 10] has been used for the voice recognition system. Speech features, sound volume and fundamental frequency are used to recognise "umm" sounds and other such sounds in order to identify pauses that frequently occurs in speech. ASIMO can estimate the number of speakers and the direction of each speaker based on the transfer function of each sound source direction. ASIMO extracts words from utterances that are closest to the ones stored in the vocabulary set. For every word in the vocabulary, there are specific tasks that ASIMO can perform.

ASIMO also has a stereo camera mounted in the head. ASIMO uses facial image recognition technology to detect a human face, face direction and face identification regardless of the person's location in the camera's field of view. ASIMO can also recognise persons facial components such as eyes, nose and mouth along with a person's facial expressions. It is also able to recognise human posture. In order to interact with humans, ASIMO uses external sensors such as, laser range finders (LRFs) that are installed in the reception area for spatial sensing, stereo cameras mounted in its head for human detection and facial features recognition and eight-channel microphone for voice recognition. The LRFs with one-dimensional scanning capability are directed at the wall surfaces, LRFs with two-dimensional scanning ability are mounted on the ceiling. The LRFs can detect visitors and their movements in the specified area.

ASIMO uses voice, image and spatial sensing information to estimate the attributes of people that are needed to deal with. Based on these attributes, different parameters such as person's interest in ASIMO, the object of person's attention and the extent of person's

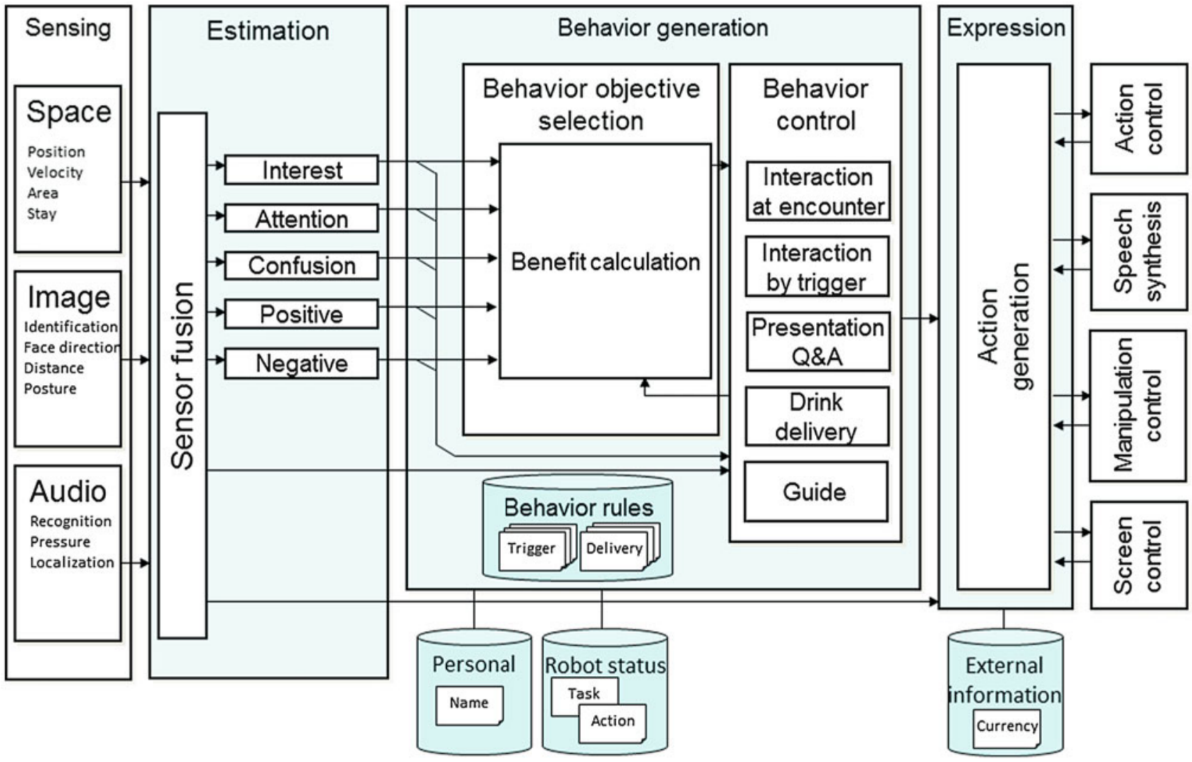


Figure 2.7: Interaction system of the ASIMO robot [Shigemi 18].

affirmation or negation are estimated using a Bayesian network to rank their likelihood. ASIMO can have different behaviour objectives depending on the number of persons it is interacting with. These behaviour objectives can be interaction at the encounter, triggered interaction, acting as a guide, giving a presentation and delivering drinks. After an optimal selection of a behaviour objective, ASIMO uses an action generator module that plans all the movements and speech of the robot for human-robot interaction. Figure 2.7 shows the interaction architecture of the ASIMO robot.

Although ASIMO is walking humanoid robot, the perception system of ASIMO is quite simple. Understanding of human nonverbal cues such as facial expressions, human postures, head gestures and hand gestures is missing. ASIMO uses speech recognition to understand verbal communication in order to react. However, behaviour analysis of human has not been considered in this robot. Asimo can adapt its behaviour by understanding speech. However, the robot is not able to analyse human behaviour over time and to adapt its behaviour based on the personality of its interlocutor.

2.2.4 SOPHIA

SOPHIA is an upper-body humanoid robot developed by Hanson Robotics [Weller 17] as shown in Figure 2.8. The robot has been presented widely in research as well as popular media. SOPHIA has natural human-like facial features through which she can exhibit more than 50 different facial expressions and emotions. SOPHIA is equipped with multiple HD cameras, mounted on its chest, for the perception of humans. Intel real sense is also mounted on its chest, and two custom 720p cameras are installed in both eyes. For speech recognition, SOPHIA uses an external microphone and also uses audio localisation array for human localisation. SOPHIA also has intelligent hands with touch sensors in the fingers.

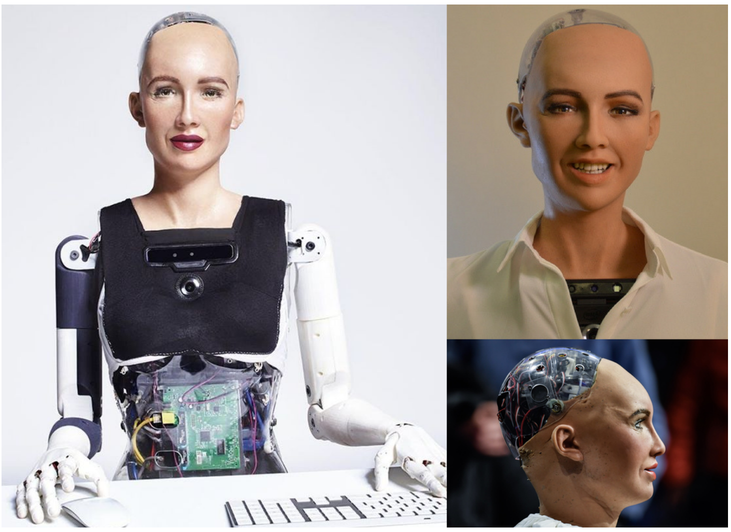


Figure 2.8: The humanoid robot, SOPHIA developed by Hanson Robotics.

SOPHIA has 83 degrees of freedom in its whole body (36 DoF in head and neck, 15 DoF in a single arm and hand, 3 DoF in the torso and 14 DoF in its mobile base). SOPHIA can recognise faces using mounted vision sensors, make eye contact, follow a person's face and recognise individuals. It uses Alphabet's google chrome voice recognition technology to recognise speech and interact with persons. She can recognise human facial expressions and can imitate as well. According to reviewers, SOPHIA can be best described as a chatbot with a face that can show several expressions [Gershgorn 17]. Due to the advanced natural language processing and artificial intelligence, it can make instant jokes and give witty replies. The robot has been demonstrated and experimented in different conferences, television shows and interviews.

Despite the focus on SOPHIA's artificial general intelligence, perception system has not been developed to understand human behaviour. SOPHIA can detect and recognise basic human nonverbal cues such as facial expressions, body postures and proxemics. However, high-level perception of human behaviour in order to express empathy and compassion has not been done. Due to the lack of analysis of human behaviour, Sophia can not adapt its behaviour based on human profile.

2.2.5 ROMAN

ROMAN is an upper-body humanoid robot developed by TU Kaiserslautern to test human-robot interaction [Berns 09], (see Figure 2.9. The robot consists of a head with an expressive face and neck, two actuated arms and torso. ROMAN can express many different facial expressions, gestures and body postures. ROMAN has 47 DoFs in his whole body which allows him to express emotions naturally. Figure 2.9 also shows the degrees of freedom of the robot. The control architecture of the robot, emotion-based architecture, is developed by Hirth [Hirth 11]. This architecture has been examined using Tangram game playing scenario [Hirth 12].

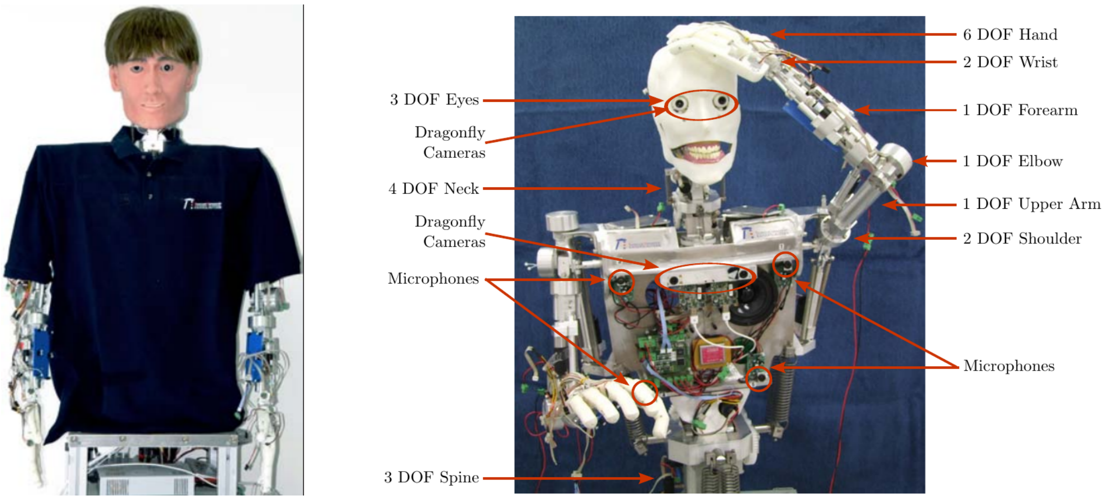


Figure 2.9: The humanoid robot, ROMAN, developed by Robotics Research Lab.

The robot has three firewire cameras that are used for perception tasks. One camera is mounted in the eyes, and it is used to detect different colours. The other camera is a stereo camera which is installed on the chest of the robot. ROMAN also has 6 microphones preamplifier units attached at different places to its body. The robot uses these microphones to localise the source of the sound, identify the speaker and recognise speech.

The perception system of ROMAN is divided into two categories: auditory perception system and visual perception system. Auditory perception system has two further branches, namely localisation and identification. The contribution of the sound source localisation is the paralinguistic aspects of sound direction and sound distance. Additionally, information of loudness is of interest to allow priority-based distinction of sound sources. Beamforming algorithm has been used for sound localization [Schmitz 09].

Speaker identification is used to identify the speaker using voice. If the person has not interacted with the robot, an initiation step is needed. The speaker identification module is divided into three parts: preprocessing, voice activity detection and voice classification [Schmitz 11]. The preprocessing step converts the input voice stream into the frequency domain. The voice activity detection generates a probability value for each microphone channel and frame of the audio stream, which is inspired by the work of Cohen [Cohen 03]. The classification module generates the current feature vectors and compares them to the set of previously trained codewords [Faust 09]. Figure 2.10 shows the auditory perception system of ROMAN.

Visual perception system of ROMAN has been developed by Schmitz [Schmitz 11]. The robot can detect human faces. It can detect human skin colour based on a combination of Gaussian probabilities which it uses to recognise different emblematic gestures used in interaction scenarios. Gestures include thumb up, thumb down, open palm, victory and loser. The emotion detector module extracts facial features from human frontal faces and according to facial action coding system (FACS), an emotional state is estimated. The

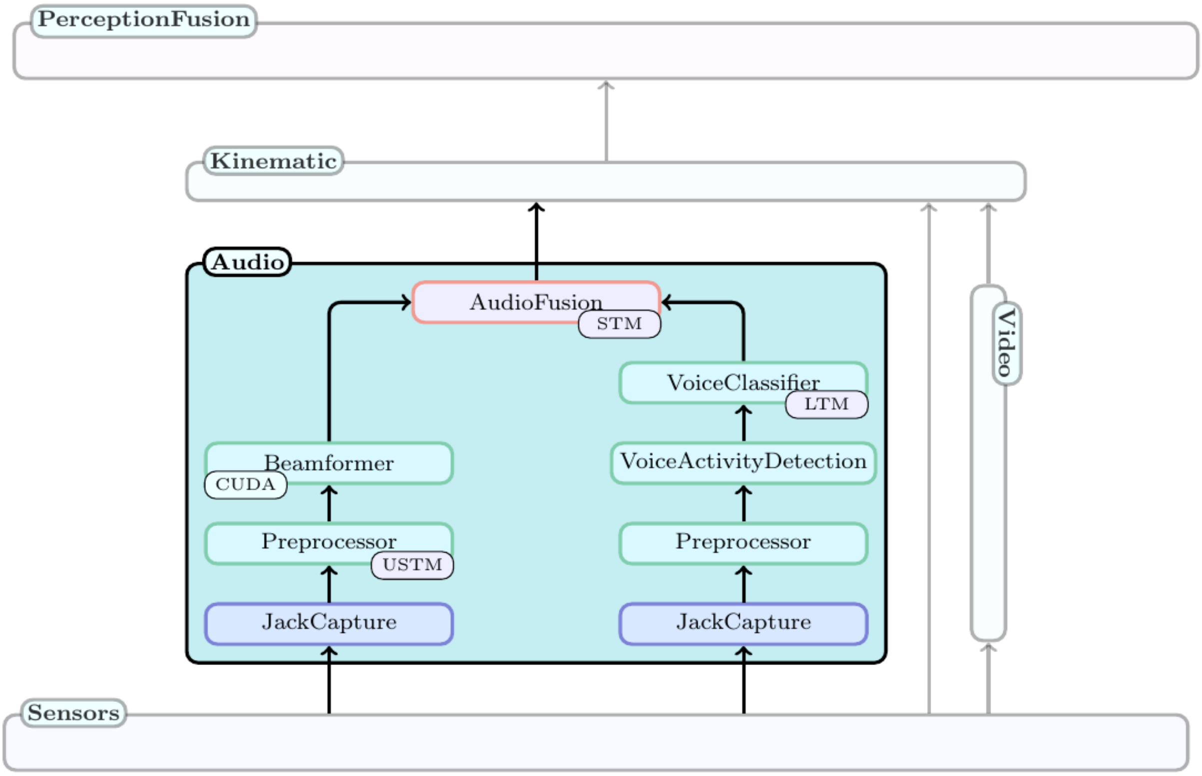


Figure 2.10: Auditory perception system of ROMAN with speaker localization and identification branches [Schmitz 11].

head pose estimator module uses the camera with a narrow field of view and estimates the head pose by matching a rigid 3D head model with a set of 2D image features. Apart from interaction partner, ROMAN also recognises the state of Tangram game. The module, Tangram detector, uses colour information to recognise Tangram board and the puzzles. Figure 2.11 shows the visual perception system of ROMAN.

The video fusion module uses all elementary perceptions like face candidates, skin colour estimation and depth image to track the head of possible interaction partners. The tracking is realised using a particle filter for each person in the environment. The multimodal fusion module combines all extracted information from visual and auditory systems into a single descriptor for each person in the environment. The kinematic module transfers the positions of all detected objects into a standard coordinate system to unify the output of the perception system. All information extracted and fused up in the multimodal fusion is then forwarded to the control architecture.

Although the perception system of the robot is able to extract several nonverbal and verbal cues, the major shortcoming is the lack of analysis of these nonverbal cues over time. Moreover, complex nonverbal cues such as detection of facial expressions, postures and body movements in real time is not possible in the robot. Nonverbal cues such as head poses, emotions, proximity and gestures are considered in this work. Despite the robot can adapt according to the external stimulus such as status of the game or the person's interest in the game scenario, the robot is not able to adapt based on human behaviour and personality.

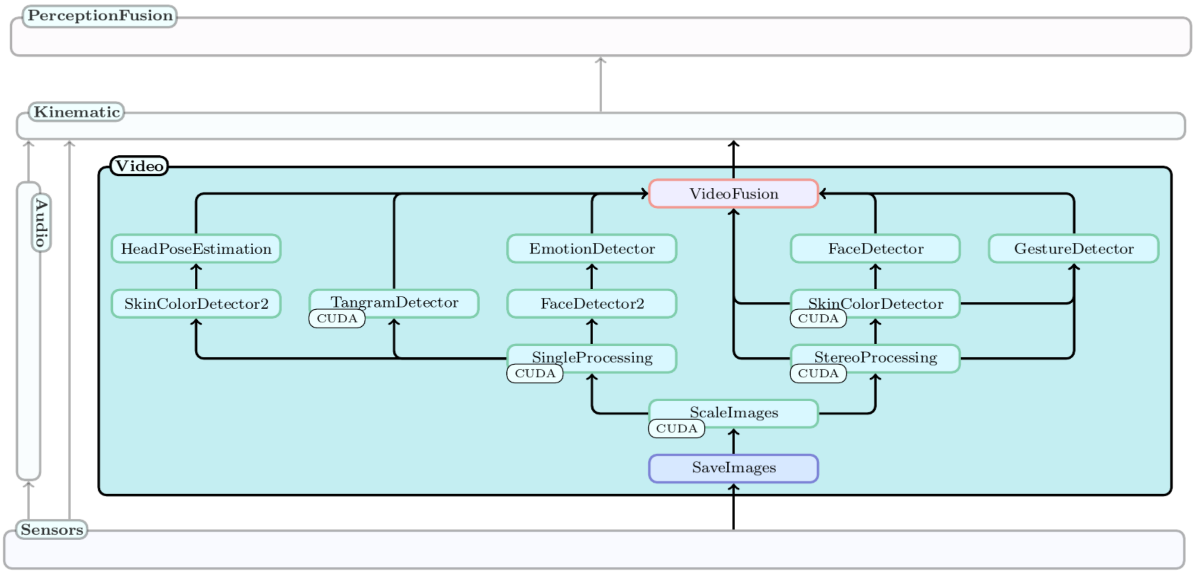


Figure 2.11: Visual perception system of ROMAN with several nonverbal cues extraction [Schmitz 11].

2.2.6 NADINE

Nadine is an upper-body female humanoid social robot that is developed at Nanyang Technological University, Singapore [Thalmann 17]. The robot face has been modelled on a Prof. Thalman to explore human-robot interaction. Nadine has a realistic human appearance with natural skin and hairs. Nadine has a total of 27 degrees of freedom for facial expressions and body movements (7 in the head, 3 in the neck, 3 in the body and 7 (x2) in the arms).

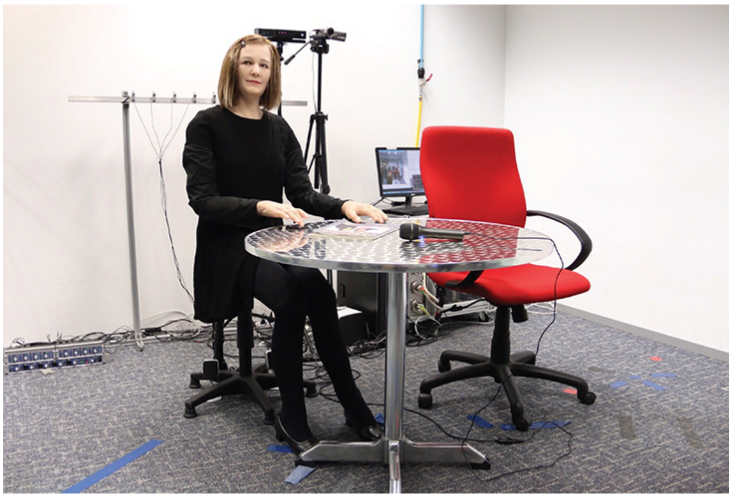


Figure 2.12: Nadine, the social humanoid robot developed by NTU Singapore [Thalmann 17].

Nadine can show emotions both in her gestures and in her face depending on the content of the interaction with the user. She can recognise people she has previously met and engage in a flowing conversation. Nadine is also fitted with a personality, meaning her mood can sour depending on what you say to her. Figure 2.12 shows the Nadine social robot. For perception tasks, Nadine uses externally mounted Microsoft Kinect sensor and microphones.

The robot architecture has been shown in Figure 2.13. The architecture consists of 3 layers, namely, perception, processing and interaction. In the perception layer, the robot perceives various stimuli which help the robot to understand the user and the environment. Visual and auditory sensors such as 3D cameras, web cameras and microphone are used as input devices to recognise user identity, position, facial emotion, actions, speech, gender and objects in the environment. These recognised stimuli are processed in the processing layer, which has various sub-modules such as dialogue processing, affective system and Nadine's memory of previous encounters with users. Finally, the verbal or non-verbal responses have to be shown on the robot using the interaction layer. The responses from the processing layer can be a head movement to maintain eye gaze, gestures and facial expressions, dialogue and tone (to show different emotions, personality).

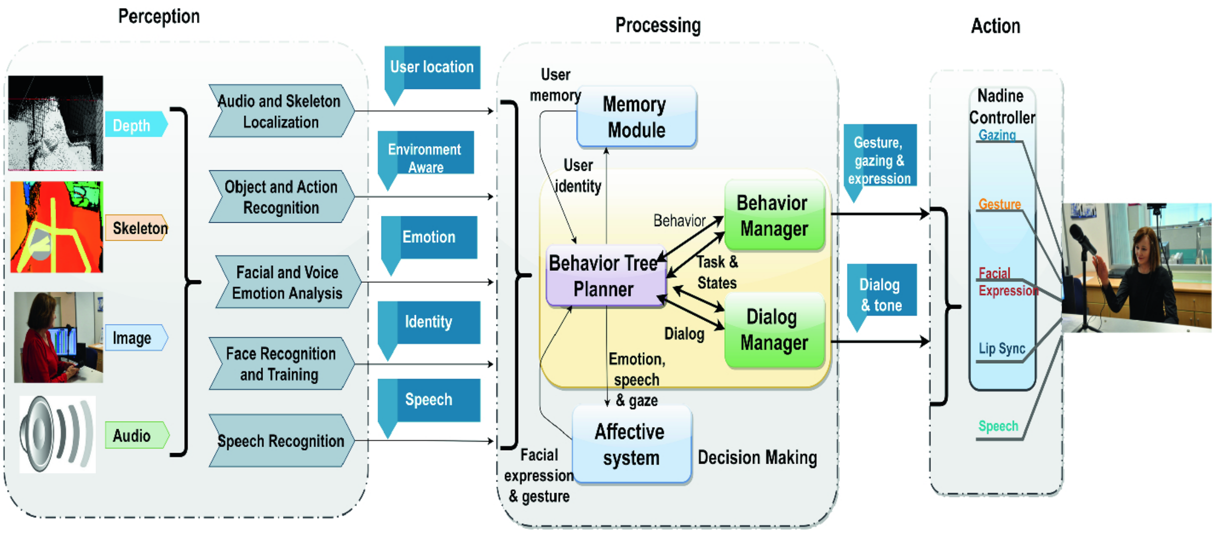


Figure 2.13: Social robot architecture of Nadine robot [Ramanathan 19].

Nadine has been used in different application scenarios for evaluation such as Nadine as a teacher, Nadine as a receptionist and Nadine as an interviewer. It can express emotions and gestures according to the perception of the user's body movements, facial expressions and speech. Despite recognising different human percepts, Nadine does not analyse human behaviour in order to express empathy and appropriate reaction during interactions.

2.3 Discussion

The ability of humans to adapt according to the behaviour of their interlocutor has been proven essential for an effective conversation in HHI. For social robots to interact naturally with humans, they must be well-adapted to human behaviour and reactions. The major objective is the adaptivity of a social robot based on the interaction partner information. This information may include an interlocutor's profile, emotions, behaviour, personality, and past interactions. Using this knowledge, a social robot should adapt accordingly [Beer 14]. An adaptive social robot is expected to have following adaptation capabilities in HRI: understand and show emotions, communication with high-level dialogue, learn/adapt according to user responses, establishing a social relationship, react according to different social situations and have varying social characteristics and roles [Fong 03].

Developing a robot with such capabilities is still an open challenge. Many of the state-of-the-art robotic systems lack in analysing human behaviour and assessment of personality traits. Most of these robots do not have sophisticated perception system that can be responsible for such complex perception tasks. They are mostly focused on the perception of speech and verbal cues to interact intelligently and adapt their behaviour. However, human nonverbal behaviour is equally or, according to many studies, more important than verbal.

Few robots have been used to assess human behaviour analysis in the literature [Salam 17] [Aly 13]. The robot used in these studies is a Nao robot [Gouaillier 09]. The problem with such a robot is its small size. The degrees of freedom in the limbs and head are extremely limited, which affect its capabilities to exhibit natural behaviours. Since the robot is not able to express human-like emotions and different complex behaviours, the interaction with these robots can not replicate human-human interaction. Moreover, because of the small size of the robot, visual sensors such as depth cameras are not installed on the robot. Instead, the sensors are generally mounted externally next to the robot which goes against the concept of anthropomorphism in social robots.

To build an intelligent social robot, the following points need to be considered.

Assessment of Human Personality Traits

Personality plays a vital role in Human-Human Interaction (HHI) as it guides the conversation towards a level of satisfaction and comfort for humans. The robot should be able to assess the personality traits of human and react accordingly. The robot should behave as an extrovert when the interlocutor is also extrovert, and sometimes robot should react as an agreeable when the interlocutor is angry and self-centred.

Psychological Background

Usually, many factors affect the assessment of human behaviours such as context, nonverbal cues, and the current situation. The interpretation of personality traits should be based on a psychological model that takes into account these factors.

This thesis contributes to the development of the perception system that can assess human personality traits for an interactive humanoid robot. The personality trait assessment system is based on psychological models of humans, which are discussed in Chaper 3.

3. Personality in Psychology

With the technological advent and constant research in the field of robotics, it is now quite practical to acknowledge the actuality of social robots being a part of human's daily life in the next decades. This necessitates and inspires the motivation of creating robots that can perceive the various learnings of life similar to humans, especially in a real-world Human-Robot Interaction (HRI). Concerning HRI, the basic expectations from a social robot are to perceive words, emotions, behaviours, and so on, in order to draw several conclusions and informed decisions for realizing natural HRI. Henceforth, assessment of human personality traits is essential to bring a sense of appeal and acceptance towards the robot during the interaction.

Human personality is made up of the characteristic patterns of thoughts, feelings, and behaviours that make a person unique. According to renowned psychoanalyst Freud [Holt 89], personality is a result of childhood experiences that are consciously or unconsciously embedded onto a person during the developmental stages. Personality plays a vital role in Human-Human Interaction (HHI) as it guides the conversation towards a level of satisfaction and comfort for humans. According to psychologists, human behaviour is known to be a combination of verbal cues together with nonverbal cues. Nonverbal cues such as temperamental characteristics are known to be innate in human beings, sometimes existing subtly or visibly. These temperaments aggregate into traits, e.g., extrovert or introvert, throughout human life through daily experiences. The significance of personality in HHI can be better exemplified by two renowned theories coming from the field of human psychology, i.e., the chameleon effect [Chartrand 99] and similarity-attraction theory [Henderson 82]. The chameleon effect explains the non-conscious human tendency to passively mimic the behaviour of one's interaction partner in a social environment.

In contrast, the similarity-attraction theory emphasizes that humans are generally attracted to and prefer the company of others who maintain morals and attitudes similar to their own. For example, it is quite often observed that there exists a sense of shared personality among friends than among random pairs of strangers. Similarity-attraction theory can be observed in people with thought processes such as not feeling alone in their belief, or the ability to predict the future behaviour of similar people in order to access the "window of bias" for enhanced relationships and validation of attraction. Besides, people tend to change their behaviour according to their interlocutor behaviour. If he/she is talkative and expressive, one also tends to be more expressive. Therefore, assessment of human personality is highly crucial for a robotic system in order to interact and adapt naturally for intelligent HRI.

The research on personality has a strong background in psychology and communication studies. A proper understanding of the technical implementation of personality traits demands decent comprehension of personality trait theories in the first place. Section 3 discusses the core theories of personality sequentially, with basic ideas, pros and cons of each theory explained briefly. The subsequent sections shed more light on some popular and universally accepted theories of personality.

Personality Trait Theories

> *"Personality is the dynamic organization within the individual of those psychophysical systems that determine his characteristics behaviour and thought."*

[Rappaport 63].

> *"The characteristics or blend of characteristics that make a person unique."*

[Weinberg 99].

The definitions presented above ensure more importance on the uniqueness of the individual, thereby adopting an ideographic point of view. This view assumes that each person has a unique psychological structure and that only one person possesses some traits; and that there are times when the process of comparing one person with other human being is next to impossible. Case studies can be applied for the purpose of information gathering in this regard.

Some of the best-known psychological theories proposed by several famous philosophers include Sigmund Freud and Erik Erikson. Specific areas of personality are the central focus of some of these theories. On the contrary, most other theories attempt to explain personality traits broadly.

Biological Theories

Approaches concerning biological aspects suggest that genetics play a vital role in personality or human behaviour. There is an interesting debate between classic nature and nurture, with the biological theories of personality siding with nature. Exhaustive research on heritability has managed to establish a link between genetics and personality traits. Studies on twins (siblings) are often used to investigate which traits might be linked to genetics versus those that might be linked to environmental variables. For example, a deeper look at the differences and similarities in the personalities of twins reared together versus those who are raised in different environments provides us with an insightful outlook in this regard.

One of the most famous theorists speaking in favour of biological factors was Hans Eysenck, who has managed to link aspects of personality to biological processes. For instance, Eysenck has argued that introverts have high cortical arousal. It leads introverts to a process that avoids stimulation. On the contrary, Eysenck also believes that extroverts have comparatively low cortical arousal, causing them to seek out stimulating experiences [Eysenck 52].

Behavioural Theories

Behavioural theories of personality are firmly of the opinion that the interaction between the individual and the environment is the key factor resulting in personality traits. The behavioural theorists extensively study observable and measurable behaviours. Behavioural theories reject the theories that keep internal thoughts and feelings into consideration. Most famous behavioural theorists include B. F. Skinner and John B. Watson, with B. F. Skinner suggesting that there is always a difference in our learning experiences. These individual learning experiences are the main reason behind our individual differences in terms of behavioural aspects.

Interestingly enough, these patterns of behaviour are learned either directly or indirectly. Learning patterns directly consider reward as positive reinforcement of good behaviour or punishment as a negative reinforcement of bad behaviour. Learning patterns through observational learning or modelling is considered to be indirect. Skinner has been a keen believer of the fact that it is simply human nature that we behave in such a way that we would receive rewards or favourable things [Skinner 53].

In order to experience reinforcement in life, we should develop positive personality traits such as those attributes included in the “agreeableness” category of the Big Five model (see Section 3.3). The attributes include being an understandable, compassionate, empathetic, and positive thinker. Taking this into account, Skinner has presented some solid arguments claiming that humans respond to every kind of reinforcement. And more importantly, our behavioural pattern and personality traits can more likely be structured and controlled by the very society we live in.

Additionally, Skinner has implied that if we want our negative traits to be changed into positive ones, we must pay attention to how we can change the environment in which we reside in the first place. This strict behaviourist point of view tries to refute other psychologists’ belief that we must alter our inner self first, i.e., our personality traits before we can fully experience the change that we want to observe in ourselves.

Psychodynamic Theories

The research works of Sigmund Freud heavily influence psychodynamic theories of personality. The theories emphasize the influence of the unconscious mind and childhood experiences on personality. The most popular and recognized psychodynamic theories include Sigmund Freud’s psychosexual stage theory [Freud 54], [Freud 01] and Erik Erikson’s stages of psychosocial development [Cherry 17].

Freud believes the three components of personality are the id, the ego, and the superego. The ‘id’ is responsible for all needs and urges, while the superego for ideals and morals. The ego moderates between the demands of the id, the superego, and reality. According to Freudian stage theory, children progress through a series of stages in which the id’s energy is focused on different erogenous zones.

Erikson also believes that personality progresses through a series of stages, with certain conflicts arising at each stage. Success in any stage depends on successfully overcoming these conflicts. Psychodynamic theories commonly hold that experiences received during childhood are responsible for shaping personality. Such theories have a remarkable association with psychoanalysis. It is a type of therapy attempting to reveal unconscious thoughts and desires of our mind. However, not all psychologists accept psychodynamic theories, and critics claim the theories lack supporting scientific data.

Humanist Theories

Psychodynamic and behaviourist explanations of personality have not been desirable to many psychologists at the time when these theories popped into the research of personality. They have felt that these theories ignored the qualities that make humans unique among animals, such as striving for self-determination and self-realization. In the 1950s, some of these psychologists began a school of psychology called ‘humanism’.

Humanistic psychologists try to see people’s lives as those people would see them. They are more likely to have an optimistic perspective on human nature [Rogers 46]. The focus is kept on the ability of human beings to think consciously and rationally. It helps to control their biological urges and to achieve their full potential. In the humanistic perspective, people are responsible for their lives and actions and have the freedom and will to change their attitudes, interests and behaviour. Two psychologists, Abraham Maslow and Carl Rogers became well known for their humanistic theories. Humanist theories emphasize the importance of a free will and individual experience in the development of personality. Humanist theorists also focus on the concept of self-actualization, which is an innate need for personal growth that motivates behaviour.

Trait Theories

The trait theory approach is one of the most prominent areas of personality psychology. According to these theories, personality is made up of many broad traits or dimensions. A trait is a relatively stable characteristic that causes an individual to behave in certain ways. Eysenck’s three-dimension theory [Eysenck 64] is one of the most famous trait theories. The five-factor theory or the Big Five theory of personality proposed by McCrae and Costa [McCrae 99] is comparatively new, but the most acceptable one in the research arena.

In Eysenck’s theory, personality questionnaires have been utilized to collect data from participants and then a statistical technique known as factor analysis is employed to analyze the results. According to Eysenck’s research, there are three significant dimensions of personality, namely extroversion, neuroticism, and psychoticism. During his initial examination, he has described only two significant dimensions of personality which he refers to as Introversion/Extroversion and Neuroticism/Stability. Extroversion and introversion provide us with information about how people like to interact with the environment in which we live in, while neuroticism and stability are more associated with emotional behaviours.

Eysenck believes that these dimensions, as a result, combine in different ways to form an individual’s unique personality. In a later stage, Eysenck included another dimension (the third dimension) known as psychoticism in the inventory. This trait is related to human attributes such as empathy, sociability and aggression. Researchers later are of the opinion that there are broadly five dimensions that constitute personalities that humans are born with [McCrae 99]. This theory is often referred to as the *Big Five* theory of personality, suggesting that there are five major personality dimensions, namely Openness, Conscientiousness, Extroversion, Agreeableness, and Neuroticism (OCEAN).

The following subsections provide us with an in-depth explanation of the most popular and effective theories concerning personality traits.

3.1 Four Temperament Theory

There is a common theory in psychology that considers the temperament to be the centre of attention. A very well known example of these theories is the *four temperament theory*. This theory introduces four basic personality types, namely sanguine, choleric, melancholic and phlegmatic. These categories are named after the bodily humours, which has a relation to our bodily fluids [Eysenck 85]. The fluids include blood, yellow bile, black bile, and phlegm.

The theory holds that the human body is filled with four basic substances, also known as humours, which are in balance when a person is in a healthy state. An excess or deficit of one of these four bodily fluids triggers diseases and disabilities. It is often assumed that diseases might also be the result of the "corruption" of one or more of the humours. Many factors including environmental circumstances, dietary changes etc. are responsible for the alteration of the humours [Lindemann 99]. Vapours inhaled or absorbed by the body have been deemed to be the cause of these deficits.

These terms used for the bodily fluids only partly correspond to the modern medical terminology, in which there is no distinction between black and yellow bile. In reality, phlegm has a very different meaning. These "humours" may have their roots in the appearance of a blood sedimentation test made in open air, which exhibits a dark clot at the bottom ("black bile"), a layer of unclotted erythrocytes ("blood"), a layer of white blood cells ("phlegm") and a layer of clear yellow serum ("yellow bile"). Figure 3.1 depicts the four elements and the transformation from one element to the other.

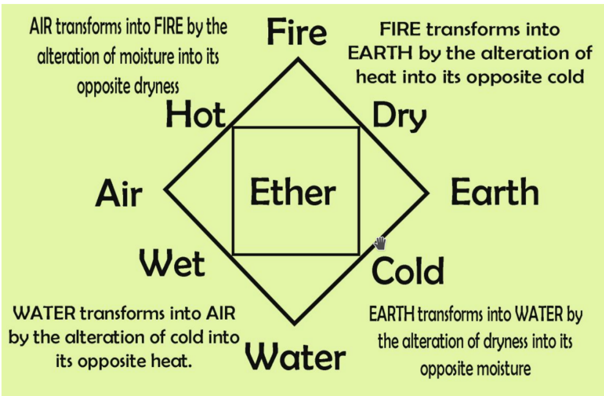


Figure 3.1: The four elements of the temperament theory [of Seville 57]

Classical melancholic temperament has a relation with the element 'earth', with the characteristic implicated being 'avoiding'. Furthermore, choleric tends to possess the element 'fire', with the characteristic implicated being 'ruling' in nature. The four temperaments are depicted in the Figure 3.2. An individual could be any combination of the following four temperaments:

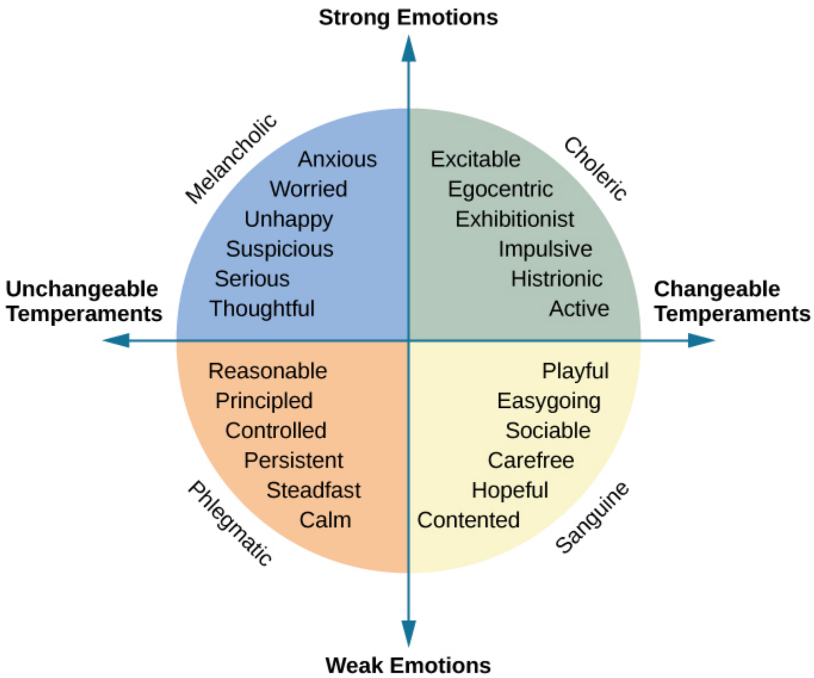


Figure 3.2: The basic concepts of the Four Temperament Theory [Spielman 18]

3.1.1 Sanguine

Sanguine personality type is mainly associated with being enthusiastic, active, and social. Sanguines tend to be more extroverted and enjoy being part of a crowd; they find that being social, outgoing, and charismatic is easy to accomplish. Individuals with this personality pass difficult times when they have nothing to be active. They are more likely to get engaged in tasks that require high risk as they find pleasure doing things at risk. As far as storytelling is concerned, the sanguine individuals often exaggerate what happened or tend to forget the main details as they focus more on making the story exciting.

3.1.2 Choleric

Choleric individuals also tend to be more extroverted. They are described as being independent, decisive, quick-thinking, active, practical, strong-willed, easily annoyed and goal-oriented. Taking responsibility of a group of people is innate in them since they are gifted with leadership qualities and ambition. Choleric personalities also have a logical and fact-based outlook on the world. They are thought to be self-confident, self-sufficient. They do not like to be verbose as they prefer direct communication. They are found out to be firm when communicating with others.

3.1.3 Melancholic

These individuals tend to be thoughtful, reserved, often anxious, analytical, detail-oriented, deep thinkers and feelers. They are introverted and try to avoid being singled out in a crowd. A melancholic personality leads to self-reliant individuals who seek perfection in their vicinity, leading to tidy and painstaking attitudes. Melancholics fear taking risks, making wrong decisions, and being viewed as incompetent. They tend to have a negative attitude toward something new. Melancholics are sceptical about almost everything, but they are creative and capable people. They tend to get bored with something once they get it figured out.

3.1.4 Phlegmatic

The phlegmatic temperament is associated with being relaxed, service-oriented, peaceful, quiet, easy-going and so on. They are more likely to have a sympathetic mind. Furthermore, they care about others, yet try not to expose their emotions. Phlegmatic individuals are also good at generalizing ideas or problems to the world and making compromises to people and ideas. Phlegmatic persons tend to live a quiet, routine life devoid of normal anxieties that people with other temperaments suffer from. They do not like to get too involved with people and life in general. It ensures that they prefer a private, home-centric and family-centric lifestyle.

3.2 Myers-Briggs Theory

There are many complex models available in psychology to recognize the personality traits of humans. Myers-Briggs theory suggests that personality types can be recognized based on our preference to particular objects, ideas, facts etc. There are four pairs in this model, namely extroversion and introversion, sensing and intuition, thinking and feeling, judgment and perception [Myers 80]. Myers-Briggs theory is an adaptation of the theory of psychological types produced by Carl Gustav Jung [Jung 71].

According to Myers-Briggs theory, an individual prefers one style more than the other for each pair. Jung also allowed a middle group in which an equal balance of the two is preferable. Combination of the letters associated with preferences is necessary to get the Myers Briggs personality type. For example, having preferences for E, S, T and J creates a personality type of ESTJ. Although a person might have preferences, he/she still uses all eight styles. It is analogous to the fact that most people are right-handed, but they still prefer to utilize both hands to accomplish tasks.

3.2.1 Extroversion and Introversion

The first pair of styles are concerned with the direction of the energy a person possesses. If there is a preference to direct the energy to deal with people, things, situations, or "the outer world", the preference is more likely for extroversion trait. On the contrary, if someone would like to direct the energy to deal with ideas, information, explanations, beliefs, or "the inner world", the preference that the person has is most probably for introversion.

3.2.2 Sensing and Intuition

The second pair concerns the type of information/things that human beings process. If there is a preference to deal with facts, something is already known, to have clarity, or to describe what we see, there is a high chance that the preference is for Sensing. If a person prefers to deal with ideas, look into the unknown, to generate new possibilities or to anticipate what isn't obvious, the preference is deemed to be for Intuition. The letter 'N' is used for intuition because 'I' has already been allocated to Introversion.

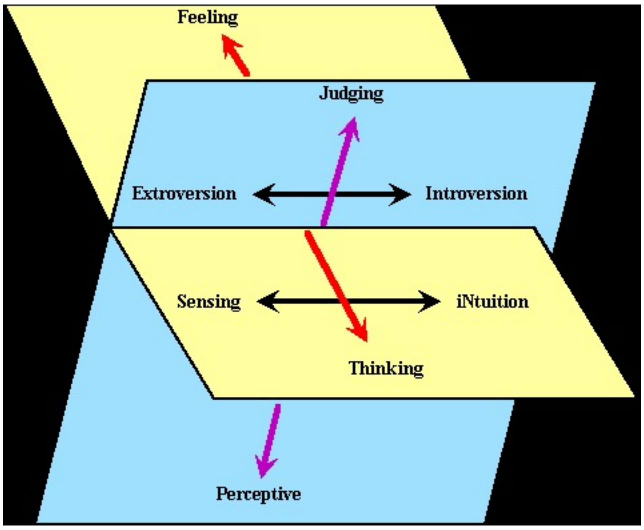


Figure 3.3: The conceptualization of the Myers-Briggs Model. (Picture taken from `http://www.nwlink.com/~donclark/hrd/styles/jung.html`)

3.2.3 Thinking and Feeling

The pair consisting of 'thinking' and 'feeling' reflects the way humans execute the decision-making process. If there is a preference to decide on the basis of objective logic, the preference is more likely for Thinking. The preference uses approaches dealing with analysis and detachment. Conversely, if there is a preference in the decision process, with values getting a higher priority, the preference is said to be for Feeling. This part of the part depends on what type of things we believe and the type of persons we feel trustworthy.

3.2.4 Judgement and Perception

The final pair of this model describes the type of lifestyle humans would like to adopt. If there is a preference that life is to be planned and well-structured, the preference is for Judging. It is not to be confused with 'Judgemental', which conceptually differs from the first part of the pair. If a person prefers to go with the flow, to maintain flexibility and respond to things as they arise, there is a high chance that the preference is for perception.

3.3 The Big Five Theory

Among all other theories presented so far in the research of personality traits, the most dominant theory has been presented by McCrae and Costa [McCrae 99], also known as the *Big Five* (BF) model. The model consists of five big dimensions, namely Open to new experience, Conscientiousness, Extroversion, Agreeableness and Neuroticism. This model is sometimes identified with the useful acronym OCEAN. As far as the *Big Five* or the OCEAN model is concerned, extroversion is more associated with activities and expressiveness. Agreeableness is more likely to have an association with being lenient, trusting, soft-hearted, generous etc. Furthermore, neuroticism has a close co-relation with temperament, self-pity, emotion, vulnerability, self-touching behaviours, depressed mood etc.

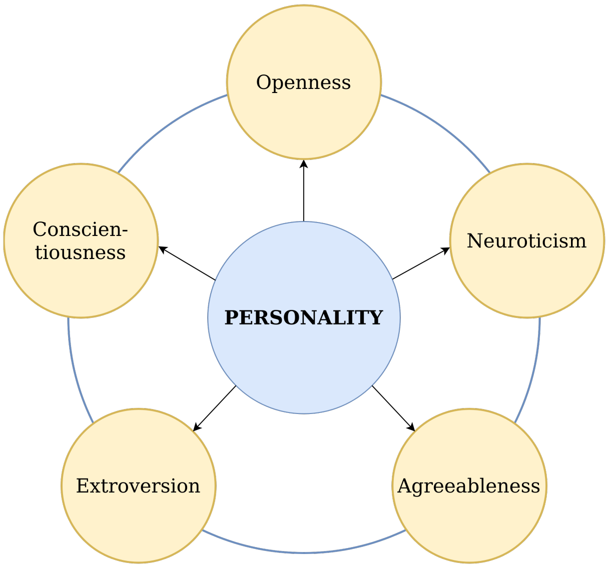


Figure 3.4: Conceptualization of the Big Five (OCEAN) model.

Persons with vivid imagination, creativity, curiosity, liberalism etc. are often deemed to be open to new experiences. Finally, conscientiousness is associated with punctuality, ambition, hard-working mentality etc. As far as this model is concerned, each dimension is considered to be a continuum or spectrum, in which the extremes are quite distinct. In other words, a person is placed somewhere on the continuum of each dimension based on the individual scores. Verbal and nonverbal facets are taken into account to recognize possible personality traits. Interestingly enough, a person can be placed in more than one dimension, but there is a dominant personality trait ingrained in each person [Jensen 16]. The following sections offer an insight into each of the dimensions of the *Big Five* (BF) model presented in order of the acronym 'OCEAN'.

3.3.1 Openness-Traditionalist

Openness is associated with an appreciation for art, emotion, adventure, unusual ideas, curiosity, and variety of experience (see Figure 3.5). This trait reflects the degree of intellectual curiosity, creativity and a preference for novelty and variety a person has [McCrae 96]. It is also described as the extent to what level of imaginative power a person possesses or how independent a person is. It depicts a personal choice and inclination for a variety of activities over a strict and inflexible routine. Unpredictability or lack of focus is often the result of high openness. However, persons belonging to this trait are more likely to engage in risky behaviour or drug taking [Ambridge 14]. Apart from that, individuals with high openness tend to lean towards being artists or writers in regards to being creative. Significance of the intellectual and artistic pursuits are critical to them [Friedman 16].

Moreover, individuals with high openness are said to pursue self-actualization specifically by seeking out intense, euphoric experiences. There is still some disagreement as to how to interpret and contextualize the openness factor. In short, Openness dimension is associated with creativity, imagination, curiosity, liberal thinking etc. Humans can deduce these facets based on empirical knowledge about an individual. From a technical point of view,

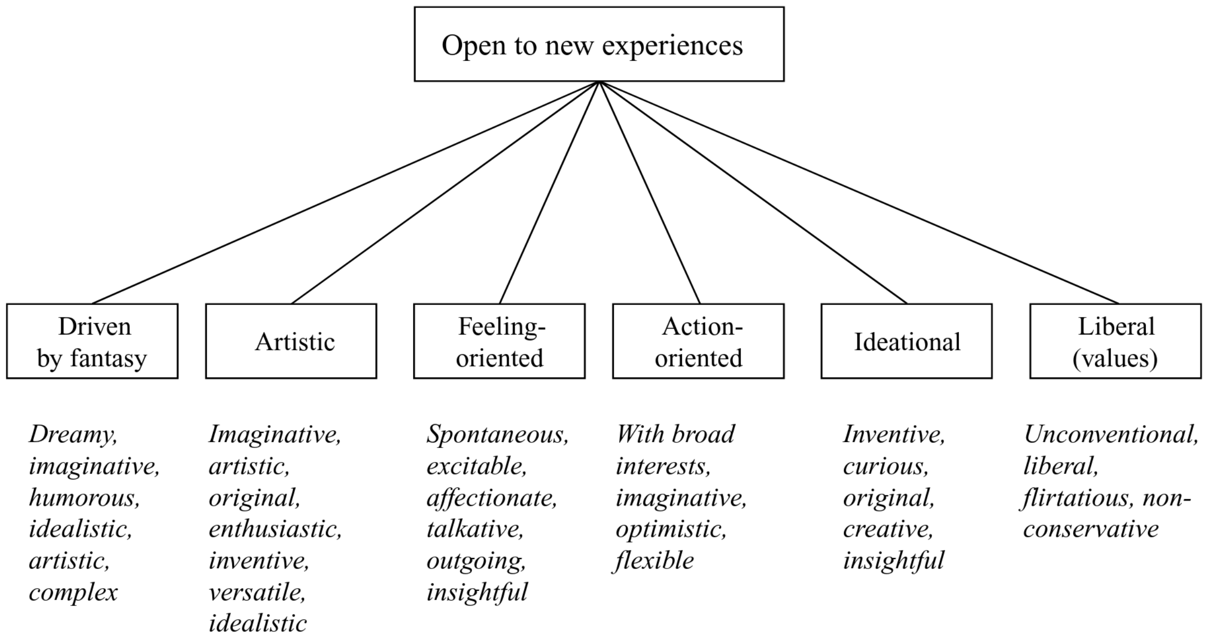


Figure 3.5: Facets of Openness personality type based on the findings of [Costa 92].

it is challenging to implement this dimension, considering only some instant nonverbal cues. The nonverbal cues relevant to openness dimension include open arms and proximity. Apart from the nonverbal facets exhibited by people belonging to openness dimension, contextual, background and cultural information is essential in this regard.

On the contrary, those with low openness seek to gain fulfilment through perseverance and are characterized as pragmatic and data-driven. They are often even perceived to be dogmatic and closed-minded. Due to their inclination towards old cultures and customs, even though some of these customs go against our conscience, they are often termed to be traditionalists. There is a marginal lack of creativity in them as compared to their counterparts belonging to the openness dimension. Traditionalism is often associated with humbleness and ‘down to earth’. Although there is a clear indication that traditionalists are conservative and incurious by nature, these individuals prefer routines to perform their day-to-day activities [McCrae 96]. Furthermore, traditionalists tend to have confidence in the established systems prevalent in society. They are more likely to protest a change that overturns the existing systems.

3.3.2 Conscientiousness-Careless

In general, conscientiousness refers to a tendency to exhibit self-discipline, act dutifully, and there is a constant strive to achieve goals that many consider being extremely difficult (see Figure 3.6). This trait has a close relation to how people control, regulate, balance and drive the impulses towards the desired goals. People with a high score on this dimension tend to mention a preference for an organized and systematic approach to performing a task than a haphazard approach [Costa 85]. Interestingly enough, the average level of conscientiousness rises among young adults and then declines among older adults. High conscientiousness is often perceived as stubbornness and obsession. Low conscientiousness is associated with flexibility and spontaneity, but can also appear as sloppiness and lack of reliability [Toegel 12]. This factor has been linked to achievement, conformity, and seeking out security, as well as relating negatively to placing a premium on stimulation and excitement [Roccas 02].

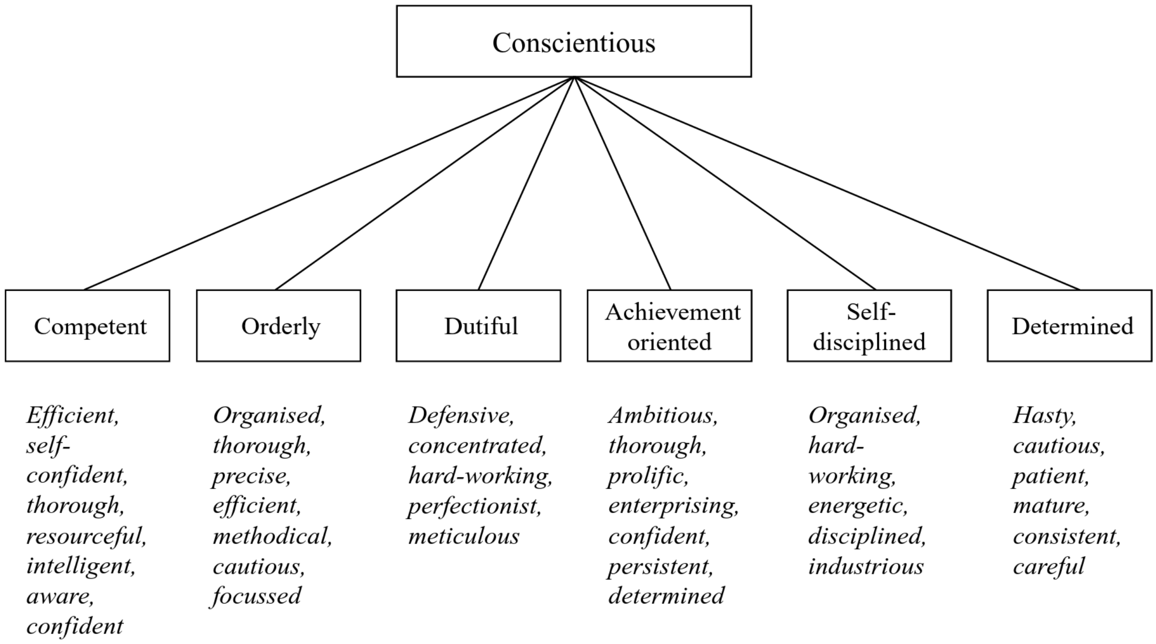


Figure 3.6: Facets of Conscientiousness personality type based on the findings of [Costa 92].

Persons scoring high on conscientiousness are also likely to value order, duty, achievement, and self-discipline, and consciously practice deliberation and work towards increased competence. Research in psychology has shown that conscientiousness also has a strong correlation to learning occurred after training [Woods 16], effective job performance [Barrick 91], and intrinsic and extrinsic career success [Judge 99]. A study conducted by Soldz and Vaillant [Soldz 99] found that conscientiousness is positively correlated with adjustment to life's challenges and the maturity of one's defensive responses. There is an indication that those scoring high on conscientiousness are often well-prepared to tackle any hindrance coming their way. This factor is also negatively correlated with depression, smoking, substance abuse, and engagement in psychiatric treatment. Conscientiousness was found to correlate somewhat negatively with neuroticism and somewhat positively with agreeableness. However, high ambition is an indispensable part and parcel of this trait [Ones 96].

People with a high score on being devoid of awareness in the day-to-day life or their overall behaviour are often considered to belong to the carelessness dimension. Carelessness behaviour can result in unintentional but severe consequences that become the primary cause of their discomfort. The effect resulting out of carelessness is often undesirable and tend to be mistaken [Pedro 11a]. The degree of lack of concern or indifference for the consequences of the action due to lack of attention may devour the origin of carelessness [White 61]. It is an hypothesis that carelessness has a close correlation with the possible cause of accident-proneness. It is observed in educational institutions that careless mistakes are committed in areas in which students already have some training. Careless mistakes are common occurrences for students both within and outside the learning environment. A lapse in judgement, also known as 'mind slips' is something responsible for the consequences careless people suffer from. There is no such conclusive evidence that provides us with the reason why careless people can not avoid making mistakes, although they are often aware of the consequences. Neurological disorders have often been blamed for the failure of the careless persons in terms of avoiding mistakes [Pedro 11b]. Carelessness is also closely associated with being lazy, aimless and late.

3.3.3 Extroversion-Introversion

Extroversion is a state in which an individual mainly obtain gratification from his or her surroundings. Extroverts feel comfortable when they interact with other people. They tend to be enthusiastic, little more talkative, assertive, and gregarious. Extroverts are energized and thrive off being around other people. They prefer to get involved in large social gatherings, such as parties, community activities, public demonstrations, and business or political groups. The nonverbal facets in which extroverts score high include duration of speech, human activeness, eye contact, proximity, etc. [Jensen 16]. Figure 3.7 shows different facets of extroverts.

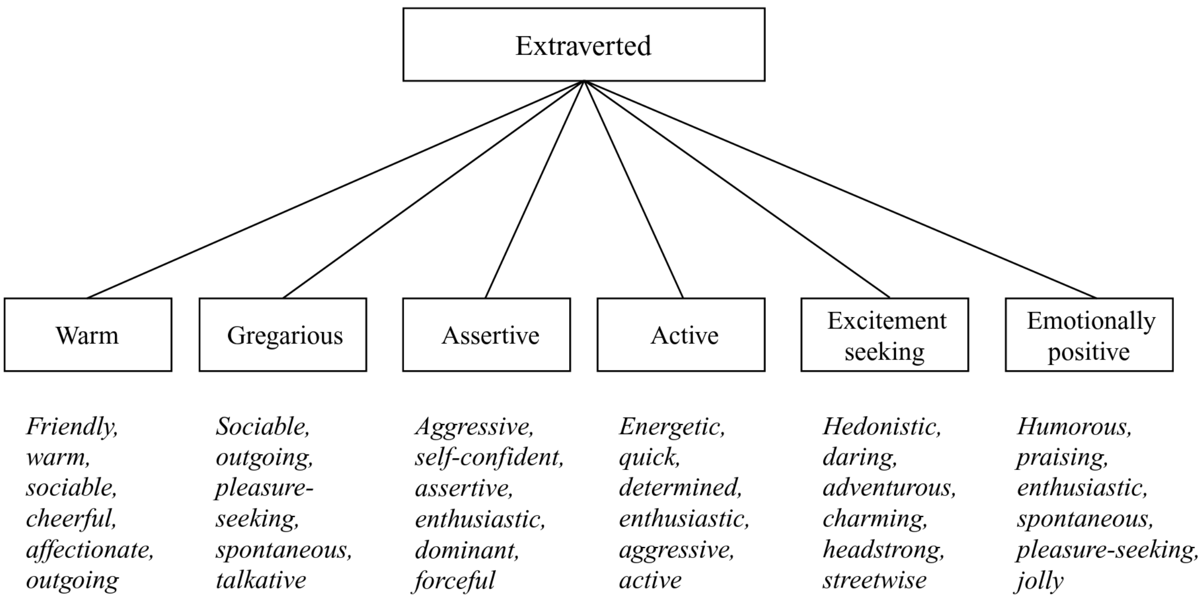


Figure 3.7: Facets of extroversion personality type based on the findings of [Costa 92].

Furthermore, extroverts also tend to work more cooperatively. An extroverted person is likely to enjoy the time they spend with people and find less reward in life if they are spending their time alone. These individuals tend to be energized when they are around other people, and they are more likely to feel bored when they are alone. Extroversion is characterized by the range of tasks performed, a high level of positive affect from external activity or situations, and the creation of vigour from external means [Laney 02]. The degree of engagement with the external world is a key factor in this trait. Extroverts enjoy interacting with people and are often perceived as full of energy. Enthusiasm is ingrained in them, with action-oriented tasks being their preference. They possess high group visibility, like to talk, and assert themselves [Olakitan 11]. Some examples of extrovert and introvert persons are shown in Figure 3.8.

Figure 3.8: Examples of extroversion-introversion trait. First two images show typical extroversion nonverbal cues, whereas the last two show signs of introversion

Introversion refers to an individual's intense interest in the matters that relate to him or her. In other words, introverts are highly interested in his mental self, or within his boundary. Introverts are typically perceived as more reserved or a loner to some extent [Myers 88]. Research in psychology and communication studies have often characterized introverts as people whose energy tends to expand through reflection and dwindle during interaction. In most of the cases, they are physically passive, quiet, sober or unfeeling. It is similar to Jung's view, although he focused on mental energy rather than physical energy. Few modern conceptions make this distinction clear.

Introverts often take pleasure in solitary activities such as reading, writing, using computers, hiking and fishing. Research has managed to find out that most of the artists, writers, sculptors, scientists, engineers, composers and inventors are all highly introverted [MacKinnon 62]. An introvert is likely to enjoy time spent alone and find less reward in time spent with large groups of people, though they may enjoy interactions with intimate friends. They are more analytical or calculative before they speak. Introverts are easily overwhelmed by too much stimulation from social gatherings and engagement. Introverts have lower social engagement and energy levels than extroverts. They tend to seem quiet, low-key, deliberate, and less involved in the social world. Their lack of social involvement should not be interpreted as shyness or depression; instead, they are more independent of their social world than extroverts. Introverts need less stimulation than extroverts and more time alone. It does not mean that they are unfriendly, antisocial or unsocial; instead, they are reserved in many social situations [Rothmann 03].

The extroversion-introversion trait is undoubtedly the central dimension of human personality theories, with the terms introversion and extroversion popularized by Carl Jung. In many research, extroversion-introversion dimension has been considered to be a single continuum (from high extrovert to high introvert): high score in one indicates having a low score in the other. For example, extroverts score high on expressiveness. Therefore, introverts have a low score on the same facet. Generally, some people have a personality trait, which is a combination of extroversion and introversion. Personality psychologist Eysenck suggests that these traits are connected somehow to our central nervous system [Friedman 16].

3.3.4 Agreeableness-Self-centred

Generally speaking, agreeableness is a person's tendency to be compassionate and cooperative toward others. This trait reflects individual differences in general and concern for social harmony. Figure 3.9 shows different facets of agreeable persons. Persons belonging to this trait show their utmost desire to keep the relationship alive, as these individuals value getting along with others. They are generally considerate, kind, generous, trusting, trustworthy, and helpful individuals. One of the most important features of agreeableness trait is that they dislike confrontation and are perfectly willing to compromise or to deny their benefits to get along with others. In their opinion, there is no need for pretence or manipulation when it comes to dealing with others. This fact ensures that they tend to be candid, frank, and genuine [Rothmann 03].

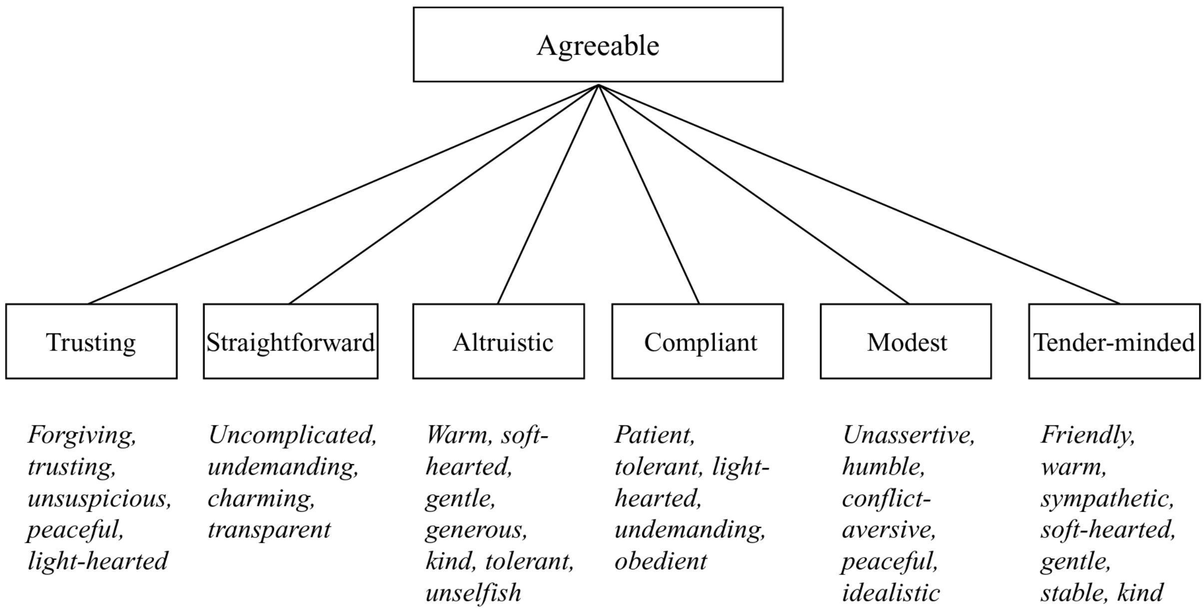


Figure 3.9: Facets of agreeableness personality type based on the findings of [Costa 92].

Needless to say, agreeableness is associated with tender-heartedness and compassion. More to the point, agreeable people also have an optimistic view of human nature. They assume that most people are fundamentally fair, honest, and have good intentions. They take people at face value and are willing to forgive and forget. Because agreeableness is a social trait, research has shown that one's agreeableness positively correlates with the quality of relationships with one's team members. Agreeableness also positively predicts transformational leadership skills. However, this dimension has also been found to be negatively related to transactional leadership in the military. A study of asian military units showed leaders with a high level of agreeableness to be more likely to receive a low rating for transformational leadership skills [Lim 04]. Some examples of agreeable and self-centred persons are shown in Figure 3.10.

Persons scoring high on self-centred behaviours place self-interest above the interest of getting along with others. They are generally unconcerned with others' well-being and are less likely to extend their helping hands for other people. It is often observed that their scepticism about others' motives causes suspicion, unfriendliness, and uncooperativeness in

Figure 3.10: Examples of agreeableness-self-centred trait. The first two images exhibit typical signs of agreeableness, whereas the last two images depict typical self-centred behaviour.

their behaviour [Bartneck 07]. They are found out to be stingy and antagonistic. It is no wonder why self-centredness is typically viewed as the most unattractive personality trait in a potential friend or partner. Most of us struggle to maintain a sense of compassion and understanding toward others.

Conversely, self-centred people, do not bother to take the time to understand another person's point-of-view or feelings. There are various degrees of being self-centred, but the general traits follow a normal structure. For example, putting themselves first, only caring about their needs and desires, being unable to see other's problems or issues, being uncaring of others. As far as their nonverbal communication pattern is concerned, they tend to exhibit fist, shake the head, pointing left or right during an interaction with other humans. These individuals like to look up, and there is a sign of angry facial expressions associated with the features discussed.

3.3.5 Neuroticism-Emotionally stable

Neuroticism is associated with experiencing more negative emotions, such as anger, anxiety, depression and fear [Jeronimus 14] (see Figure 3.11). Often termed as emotional instability, ensuring a high range of variability in the behavioural pattern. If it is reversed, it is known as emotional stability. According to Eysenck's theory of personality [Eysenck 67], there is a low tolerance for stress or aversive stimuli in neuroticism. It is a classical temperament trait that has also been studied in temperament research for a long period. Later, the concepts were adjusted and adapted by the *Big Five* model [Kagan 09] of personality. Since the main properties of temperament traits are stability in a lifetime and its neuro-physiological basis, researchers used these properties of neuroticism to support *big five* model.

Those who score high in neuroticism are emotionally reactive and vulnerable to stress. On top of that, they also tend to be flippant in the way they try to expose themselves. There are high chances that they would treat ordinary situations as threatening and counterproductive. They often interpret minor failure as extremely difficult to overcome. The adverse emotional reactions they experience tend to persist for unexpectedly long periods. In other words, they are often in a lousy temperament, and they tend to lose mental rigidness (see Figure 3.12). Studies show that neuroticism is connected to a pessimistic approach toward work, confidence that work impedes personal relationships, and apparent anxiety linked with work [Fiske 10].

Furthermore, individuals scoring high on neuroticism may show more skin-conductance reactivity than those scoring low on neuroticism [Norris 07]. These problems in emotional regulation can diminish the ability of a person scoring high on neuroticism to think

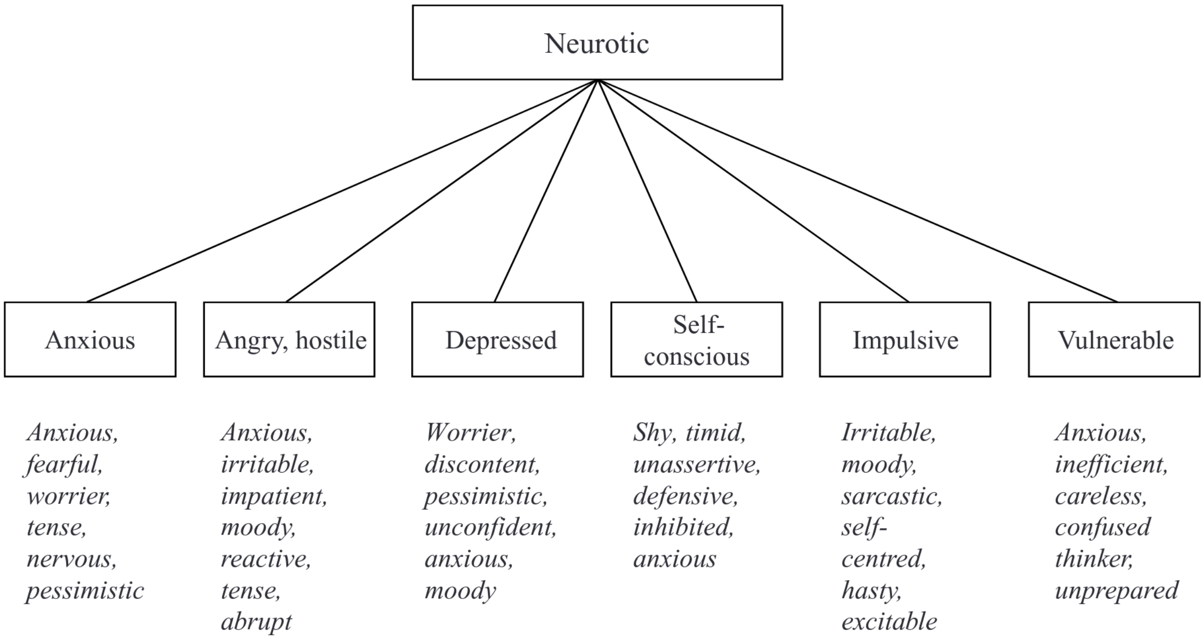


Figure 3.11: Facets of neuroticism personality type based on the findings of [Costa 92].

clearly, make decisions, and cope effectively with stress. Lacking contentment in one's life achievements can correlate with high neuroticism scores and increase one's likelihood of falling into clinical depression [Charles 08]. Although individuals high in neuroticism are more likely to experience adverse life events, neuroticism also changes with respect to positive and negative life experiences [Jeronimus 14]. Neuroticism is similar but not identical to being neurotic in the Freudian sense. Therefore, some psychologists believe in calling neuroticism by the term ´emotional instability' to differentiate it from the term ´neurotic'. Self-touching behaviour is prevalent in this dimension. They like to look down when engaged in an interaction with others.

Figure 3.12: Examples of neurotic behaviour in terms of nonverbal cues.

The inverted side of neuroticism is emotional stability. Individuals who score low in neuroticism are not easily upset and are less emotionally reactive. They tend to refrain from showing emotional variability. Instead, they are more likely to be calm, emotionally stable, and free from persistent negative feelings or experiences in life. Just because there is a scarcity of negative feelings does not mean that low scorers experience a lot of positive feelings all the time [Dolan 06]. Emotionally stable persons are more likely to be calm, quiet and even-tempered. They are satisfied with the achievements or failures of their lives.

Apart from this, they are found out to be unemotional and comfortable in response to positive or negative life experiences. As far as the nonverbal facet is concerned, emotional stability is associated with the exhibition of comparatively less number of facial expressions during conversations or social interactions [Jensen 16].

3.4 Temperament Framework for Personality Traits Assessment

A fundamental need of any unified science is the existence of a few basic dimensions that can be used for analysis and representation purposes. The fundamental difference between natural and social sciences is that the natural sciences have these dimensions, e.g., length, time and mass, while social sciences have lacked to define such representative dimensions. However, since the mid of the last century, many researchers have tried to explain human emotional states in multiple dimensions. However, the definition of emotion varies for each researcher, who adopted one or more dimensions to define it.

According to Wundt and Judd [Wundt 97], the three dimensions of emotions are namely, “pleasurable vs unpleasurable”, “arousing vs subduing” and “strain vs relaxation”. Many emotional spaces have been presented in psychology. Among them, the prominent ones are the circumplex model by Russell [Russell 80] and the Positive Activation-Negative Activation (PANA) model presented by Watson and Tellegen [Watson 85]. The circumplex model of affect suggests the distribution of emotion over a circular two-dimensional space, which consists of valence dimension on the x-axis and arousal on the y-axis. On the other side, the PANA model, also known as the consensual model, is known to be a 45-degree rotational version of the circumplex model. In PANA, the dimensions of arousal and valence lay at an angle of 45-degrees to the x-axis and y-axis, represented by negative activation and positive activation, respectively.

3.4.1 Pleasure, Arousal, and Dominance Emotional Space

There exists another renowned three-dimensional emotional space, called Pleasure-Arousal-Dominance (P.A.D.) emotional space, presented by Russell and Mehrabian [Russell 77]. The P.A.D. model aims to describe and measure emotional traits that correspond to human personality traits. The three dimensions are defined to be bipolar such that pleasure is described as a continuum that ranges from intense pain or unhappiness on one end to intense happiness or ecstasy on the other. Arousal has been reported to range from sleepiness and drowsiness to a high level of alertness and excitement. Dominance varies from emotions of a complete absence of control or impact over events to feeling influential and in control of the situation at the opposite extreme.

With regard to human emotion, researchers have categorized emotions as either discrete and different or grouped on the basis of dimensions. In this dimensional approach, emotion serves as a point in the continuous emotional space represented by distinct dimensions that strive to realize human emotions. The model also attempts to draw a parallel with the interconnection of different emotional states based on common neural systems.

If emotions are described appropriately in terms of pleasure-displeasure, arousal-nonarousal, and dominance-submissiveness, then, identification of fundamental dimensions of temperament follows simply and logically [Mehrabian 78]. Temperaments can be defined as an individual's generalized emotional predisposition and be assessed in terms of characteristic patterns and/or averages of the states of pleasure, arousal, and dominance across typical life situations. This framework is based on pleasure, arousal, and dominance (P.A.D.) emotional space. After extensive research, authors have defined these three domains as follows. Pleasure can be determined using cognitive judgments of evaluation, i.e., higher evaluations of stimuli associated with greater pleasure induced by stimuli. Arousal corresponds to judgments of high-low stimulus activity using measure of stimulus "information rate". Dominance is defined as judgment of stimulus potency, with more significant the influence of stimuli corresponding to lower values of dominance.

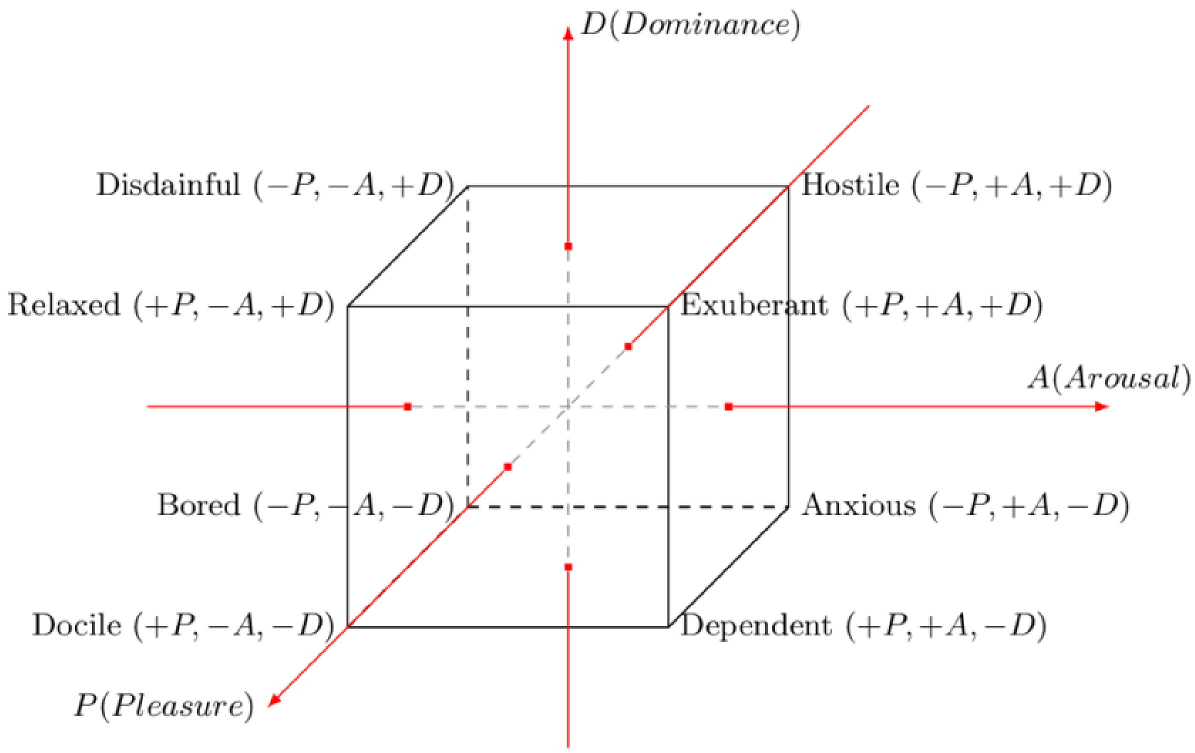


Figure 3.13: Pleasure, Arousal, and Dominance Emotional Space (figure used from [Tarasenko 10].

3.4.2 Temperament Framework

Trait Pleasure, Trait Arousability, and Trait Dominance define three orthogonal almost independent axes of temperament space, as shown in Figure 3.13. Points in this space define individuals, segments or regions of the space define personality types, and straight lines drawn through the intersection point of three axes define various personality dimensions. According to the study [Mehrabian 96], P.A.D. emotional space can be divided into 8 regions based on both the extremes of each axis, denoted as +P and -P for pleasant and unpleasant, +A and -A for arousable and unarousable, and +D and -D for dominant and submissive, temperament, respectively. The octants of temperament space are as follows:

- Exuberant $(+P + A + D)$ vs. Bored $(-P - A - D)$

- Dependent $(+P + A - D)$ vs. Disdainful $(-P - A + D)$
- Relaxed $(+P - A + D)$ vs. Anxious $(-P + A - D)$
- Docile $(+P - A - D)$ vs. Hostile $(-P + A + D)$

Using the three dimensions, pleasure, arousal, and dominance, Mehrabian has formulated 59 individual measures that correspond to human personality traits. It has been demonstrated that traits are symmetrically related to one another based upon the P.A.D. dimensions. Some of these equations are shown in equation 3.1. All the variables in the equations must be standardized. In each equation, a given personality scale is described as a function of a linear combination of Trait Pleasure (P), Trait Arousability (A), and Trait Dominance (D) scale scores.

$$\begin{aligned}
Intellect &= 0.14P + 0.20A + 0.48D \\
Achievement &= 0.13P + 0.60D \\
Extroversion &= 0.21P + 0.17A + 0.50D \\
SocialDesirablility &= 0.34P - 0.26A + 0.17D \\
ArousalSeeking &= 0.14P + 0.26A + 0.55D \\
Aggression &= -0.36P + 0.20A + 0.28D \\
TraitDominance &= 0.72D \\
PhysicallyActive &= 0.26P + 0.40D \\
Anxiety &= 0.24A - 0.20D \\
Shyness &= -0.29P + 0.13A - 0.56D \\
SensitivitytoRejection &= 0.14A - 0.71D \\
Nurturance &= 0.41P + 0.12A + 0.17D
\end{aligned} \tag{3.1}$$

The above mentioned measures are defined with the help of pleasure, arousal and dominance emotional space. Consider an example of arousal seeking trait. It can be defined as, a person that looks for excitement, change, new environments, taking the risk, etc. This personality scale is related to exuberant temperament, but instead of characterizing specific interpersonal orientations, this trait tends to characterize ways in which individuals generally relate to situations. People that seek change, risk, new environments, and unusual stimuli come under this trait. This trait, therefore, is correlated strongly to dominance and arousal. Similarly, sensitivity to rejection trait defines a person who is affected easily by the negative remarks of others. According to [Mehrabian 96], sensitivity to rejection is simply a general measure of social submissiveness and understandably, a strong negative correlate of trait dominance. The 12 mentioned traits are discussed in Chapter 6 in detail.

Computed personality scores, using this framework, are useful when a subject's scores on several personality dimensions are needed. It can be the case during human-robot interaction when an individual interacts with a robot and exhibit mentioned personality traits. Subjects are administered on the P.A.D. scales during an interaction. Once the P.A.D. scores are available, other traits scores can be computed in real-time.

4. Concept of Intelligent Human-Robot Interaction

Humanoid robots, built to resemble human body (e.g., Atlas robot by Boston Dynamics built for search and rescue tasks or Sophia robot by Hanson Robotics built to communicate and interact with humans), will soon become part of our daily-life activities in coming years due to the rapid advent of high-end computer systems and technology.

Providing humanoid robots with social skills, e.g., understanding human emotions and intentions, has been studied for many years and is still considered as an active research area. However, several challenges have been recorded in this field due to the complexity of human-robot interaction (HRI) and interpersonal communication. Currently, researchers are working on enabling these humanoid robots to function in a human environment for normal daily life activities, such as working side by side with humans in factories (e.g., [Sanfilippo 15] and [Yang 17]), working as receptionist in offices (e.g., [Gardecki 17] and [Gardecki 18]), assisting patients or elderly people at home (e.g., [Kobayashi 11] and [Joseph 18]), etc. To function properly in human environments, these humanoid robots need to be able to communicate and interact with human beings in order to understand and perceive human's commands, feelings, intentions, and demands. Therefore, researchers have focused on each of these mentioned areas to make better interaction between the human and the humanoid robot, e.g., [Al-Darraji 16b] and [Yang 12].

Several communication and interaction modes from verbal to non-verbal have been developed by researchers to make the collaboration between humans and these robots more intuitive. However, because of limitations of the verbal communication (such as understanding the semantics, working in a noisy environment, non-native speakers, speech disorder, etc.), non-verbal communication modes have been more popularised amongst the researchers to provide interaction between human and social robots. Naturally, human use different non-verbal communication modes (such as human postures, hand gestures, facial expressions or bodily cues) to give commands or to express their internal emotions and feelings.

Human interactions are highly dominated by the perception of social and behavioural traits. Personality is a significant part of behavioural traits that expresses the characteristics of individuals in different situations. Humans constantly assess the personality of their counterparts to interact robustly to flourish Human-Human Interaction (HHI). The old but famous saying *First impression is the last impression*, although one of many clichés, is also based on the fact that humans tend to evaluate personality and make an assumption of oneself. It can be validated from comprehensive research conducted in social psychology about the perception and evaluation of behavioural traits that involve spontaneous, unintentional and unaware processes [Uleman 08].

In Human-Robot Interaction (HRI), a direct relationship between personality and behaviour has long been recognised [Christensen 05] [Nass 01]. There are many implications of assessing personality that can be reached based on the HHI scenarios. Persons engaged in an interaction behave differently based on the personality types they possess and the overall environment in which they act. For example, extroverts seem to have more control and comfort over an interaction, whereas introverts often show less intimacy, control and dominance over a conversation. More to the point, submissiveness - dominance culture in HHI is relatively trivial and can be assessed based on the verbal and nonverbal cues. According to Nass et al. [Nass 01], generally, humans are more likely to interact with others having similar personality type. This fact can be observed from our social circle, e.g., outgoing people tend to make more interactions with people having similar type of personality. Furthermore, sexual orientation, age, social status, etc. also play a significant role in this regard. It shows the significance and role of human personality in daily life interaction scenarios. However, assessment of human personality traits is a challenging task, which, sometimes, humans find it difficult to do. In the following section, a list of challenges and problems associated with human personality traits assessment has been described.

4.1 Problems and Challenges

In the literature, there exist minimal works that are focused on assessing human personality traits in the field of human-robot interaction. Most of the state-of-the-art humanoid robots do not analyse human behaviour and behavioural traits. The perceptual ability of these social robots is quite naïve, and they only recognise low-level percepts, such as faces and simple gestures. Several research-works in the literature for assessing human personality traits are presented as a standalone system. These works do not cater to the interaction part between robot and human, rather focus on the interaction between a system and human. Therefore, challenges associated with a humanoid robot, such as *'how a robot should behave?'*, *'what expressions should it express?'*, *'when to reply or when to ask a question? (turn-taking)'* and so on, have not been explored in these works.

Few robots have been used to assess human personality in the literature [Salam 17] [Aly 13]. The robot used in these studies is a Nao robot. The problem with such a robot is its small size. The degrees of freedom in the limbs and head are extremely limited, which affects its capabilities to exhibit natural behaviours. Since the robot is not able to express human-like emotions and different complex behaviours, the interaction with these robots can not replicate human-human interaction. Moreover, because of the small size of the robot, visual sensors such as depth cameras are not installed on the robot. Instead, the sensors are generally mounted externally next to the robot which goes against the concept of anthropomorphism in social robots.

Some systems have been presented in the field of human behaviour analysis which are shown to be effective in different applications. However, there are some significant shortcomings in these systems. These systems generally work in a constraint environment. Although some of them report decent recognition rates, however, in an unstructured environment, many of these systems report poor results. The approaches presented in these works consider either facial expressions or gestures. Other nonverbal facets such as proximity, speech duration, head gestures and body movements also play a critical role in the accurate assessment of human personality traits and are not considered in most of these systems. Another critical shortcoming is the lack of real-time assessment of human personality traits based on human appearance in daily life interactive scenarios. For a social robot to show appropriate behaviour and adapt its reactions according to the personality trait of the interlocutor without any delay, the assessment of personality traits should happen in real-time. The assessment of personality traits in real-time depends upon the robust and efficient recognition of nonverbal features in real-time.

Although a limited number of technical systems have been reported in the literature for real-time personality traits assessment, these systems at best can *only* recognise the *Big Five* (BF) personality traits. They are unable to distinguish between subtle personality traits, for example, shyness and introversion or dominance and aggression. According to Watson and Clark [Watson 97], extroversion can be subdivided into the more specific facets of assertiveness, gregariousness, cheerfulness, and energy. Similarly, neuroticism can be subdivided into loneliness, anxiety, and sensitivity to rejection, while shyness is the part of introversion trait. The assessment of these subtle traits is highly relevant in the context of HRI.

4.2 Human Personality Traits Assessment

The primary goal of this thesis is to develop a system that enables a robot to synthesize an appropriate behaviour adapted to human profile (i.e., personality). Human personality is made up of the characteristic patterns of thoughts, feelings, and behaviours that make a person unique. Personality plays a vital role in Human-Human Interaction (HHI) as it guides the conversation towards a level of satisfaction and comfort for humans. According to psychologists, human behaviour is known to be a combination of verbal cues together with nonverbal cues. Nonverbal cues, such as temperamental characteristics, are known to be innate in human beings, sometimes existing subtly or visibly. These temperaments aggregate into traits, e.g., extrovert or introvert, throughout human life through daily experiences.

4.2.1 Significance of Personality in HRI

Before building a system for personality assessment, the question *"why should personality traits be considered in human-robot interaction?"* needs to be answered. This can be done by analysing role of personality in human-human interaction (HHI). The significance of personality in HHI can be better exemplified by two renowned theories from the field of human psychology, i.e., *the chameleon effect* [Chartrand 99] and *the similarity-attraction theory* [Henderson 82]. The chameleon effect explains the non-conscious human tendency to passively mimic the behaviour of one's interaction partner in a social environment. In contrast, the similarity-attraction theory emphasizes that humans are generally attracted to and prefer the company of others who maintain morals and attitudes similar to their own. For example, it is quite often observed that there exists a sense of shared personality among friends than among random pairs of strangers. Similarity-attraction theory can be observed in people with thought processes, such as not feeling alone in their belief, or the ability to predict the future behaviour of similar people in order to access the "window of bias" for enhanced relationships and validation of attraction. Subsequently, people tend to change their behaviour according to their interlocutor's behaviour. If he/she is talkative and expressive, one also tends to be more expressive.

In human-robot interaction, the relationship between personality and behaviour has been found [Nass 01] [Woods 07]. In the context of human modelling and adapting the dialogue of a machine (i.e., a humanoid robot or a computer) to the personality of the interacting human, [Nass 01] and [Tapus 08] has proved empirically that the person interacting with a robot would spend more time if the robot personality matches with the personality of a person which validates the similarity-attraction theory in human robot interaction scenarios.

There is another compelling theory called complementary attraction which describes that individuals also attract towards those people whose personalities are complementary to their own personalities [Isbister 00]. In order to understand the effect of this theory, [Lee 06a] have used AIBO robot and conducted long-duration of experiments. The authors have found the participants have been interacting with the robot when the robot has a complementary personality than when it has a similar personality. The good example of complementary theory is the long-term relationships between two persons with particular roles such as in marriages and in office environment. Similarly, several studies have also found the relationship between human personality and the proximity behaviour. Tapus et al. [Tapus 08] have found that people with extroversion personality type are more lenient with their personal space invasion by a robot than introverts.

Hence, the assessment of human personality trait is highly critical for a robot to adapt its behaviour appropriately during human-robot interaction. Enabling a robot to synthesize an appropriate behaviour adapted to human profile in the real-world can be done by assessing human personality traits by implementing psychological theories and models on a real social humanoid robot. In order to build such a system and address the above-mentioned problems and challenges, a personality trait assessment architecture has to be modelled using psychology and cognitive studies. This architecture is also responsible to enable a robot to adapt its behaviour and behave in an appropriate manner. Figure 4.1 illustrates the personality architecture. We describe the requirements of the framework in the following sections.

4.2.2 Visual Perceptual System

The visual perceptual system of the robot is responsible for perceiving, interpreting and making sense of the world. Analysis of human behavioural traits requires a variety of perceptual skills. These perceptual abilities, which include an understanding of human nonverbal cues such as emotions, gestures, postures, and many more over time, help the robot to perceive different human actions. By using the fusion of such different perceptual abilities enables a robot to understand human emotions, behaviour and personality traits.

Several studies have emphasised the importance of nonverbal cues in human-human interaction. Understanding nonverbal cues of the interaction partner is one of many perceptual skills that humans have.

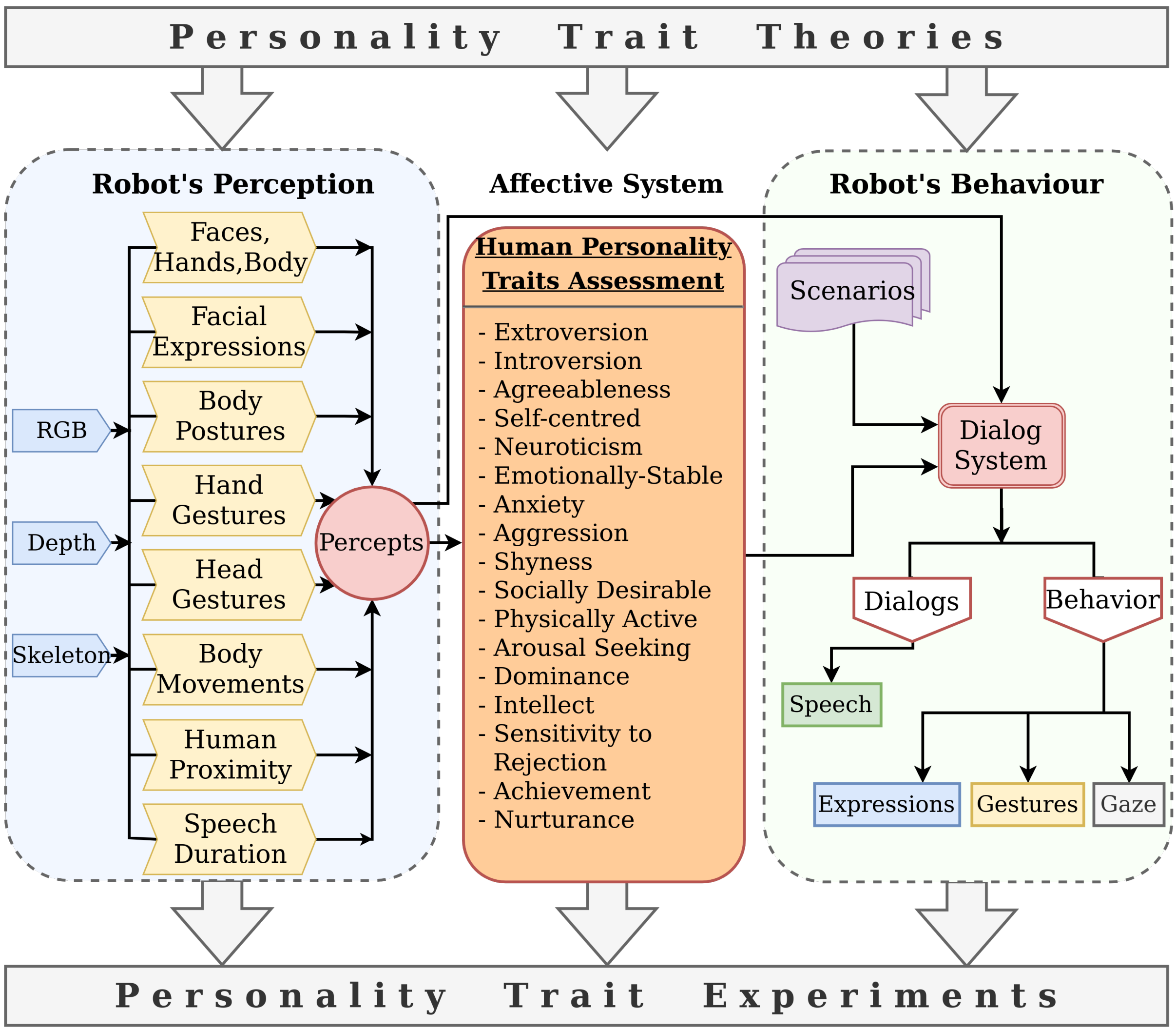


Figure 4.1: Personality architecture of a social robot. Using personality trait theories from psychology, several nonverbal facets, *percepts*, are recognized from visual sensory data. The percepts are analysed over time to predict and estimate human personality traits. Robot's control use the personality information and show empathy by verbal dialogues and physical behaviour.

Several studies have been conducted on nonverbal cues and their relation to judgments of emotions and personality (e.g., [Ekman 69], [Scherer 77]). Nonverbal cues are generally categorised into three domains of dynamic cues: face, i.e., facial expressions, body, i.e., body language, which can also be divided into gestures and postures, and tone, i.e., paralanguage. An individual who shows cues such as a huddled posture or head pointed towards the ground might, for example, be seen as someone with a (conflict) avoidance motive that wants to maintain a stable relationship with possible interaction partners. This person might then be (correctly) identified as being submissive or introverted.

Similarly, extrovert people tend to be quite active, and their excessive hand movements during conversations most often show confidence and control [Oberzaucher 08]. Similarly, nonverbal cues such as, body postures, hand gestures, facial expressions, human proxemics and speech duration, also play an important role in extracting the emotional state of an interlocutor. It makes the nonverbal cues the basic building block of the personality assessment system. Several psychologists have pointed out the various combination of nonverbal cues that denotes certain human behaviour trait. Accurate and robust detection of nonverbal cues is highly essential due to their significance in human interactions. Furthermore, the detection and recognition of these nonverbal cues needs to be validated in complex and unstructured environments to assess their reliability. In addition, recognition process should also be in real-time so that the affective system should process the percepts in a timely manner.

4.2.3 Assessment of the Big Five Personality Traits

As mentioned earlier, human interactions are highly dominated by the perception of social and behavioural traits. Personality is a significant part of behavioural traits that expresses the characteristics of individuals in different situations. The ability of humans to adapt according to the behaviour of their interlocutor has been proven essential for an effective conversation in HHI. For social robots to interact naturally with humans, they must be well-adapted to interaction partner behaviour, reactions and other personalised information. This information may include an interlocutor's profile, emotions, behaviour, personality, and past interactions. The research aim of this thesis is to enable a humanoid robot to understand and assess human behavioural and personality traits using visual sensor in daily life scenarios.

Different psychology theories of personality traits have been studied, and ultimately, the *Big Five* (BF) personality traits theory has been considered. As mentioned earlier, BF personality theory defines a person in *five* unique dimensions in which each dimension is a continuum. The research goal is to enable a humanoid robot to analyse human personality traits and adapts its behaviour. Different methodologies have been tried to assess big five personality traits, such as weighted summation of nonverbal features over time [Berns 18] and k-means clustering of nonverbal features [Zafar 18b]. However, due to inherent problems of these methods, such as correct estimation of weights and apriori specification of the number of clusters, there is a need to develop a system that can accurately assess human personalities. The affective system of the robot should process the percepts from the visual perceptual system of the robot according to their significance and assess personality traits in real-time.

Other important aspect is the validation of personality assessment system. The personalities of the subjects must be evaluated in different scenarios to measure the accuracy of the system. Scenarios must also depict a standard interaction from daily life. It does not need to be a complicated scenario. However, it must be designed in a way that the subjects can express themselves. Moreover, since not many personality databases are available, a psychology expert must also be present to validate the findings of the system.

4.2.4 Assessment of Subtle Personality Traits

As noted earlier, *big five* personality traits are generalised traits of a person. However, these traits do not distinguish a subset of *big five* categories. For example, aggression, dominance and physically active are sub-traits of self-centred and extroversion dimension. The assessment of these subtle traits is highly significant and relevant in the context of HRI. This thesis proposes to use the P.A.D. emotional space for the assessment of human personality traits using the framework presented in the literature [Mehrabian 96].

Using the three dimensions, pleasure, arousal, and dominance, as explained in Chapter 3, the author has formulated 59 individual measures that correspond to human personality traits. It has been demonstrated that traits are symmetrically related to one another based upon the P.A.D. dimensions. Although the formulated traits are of a wide range, only 12 out of 59 traits are realised in this thesis. These traits are chosen according to the experimental restrictions and based on the knowledge of nonverbal cues associated with them. Personality traits such as mysticism, loneliness, and anorexic require either verbal or contextual information or both for an accurate assessment. Furthermore, even humans find it challenging to assess these traits in human-human interaction.

In order to use the temperament framework, human emotional state is represented in pleasure, arousal and dominance emotional space. Using psychology and cognitive science studies as the reference, perceptual skills of the robot can be used to estimate the P.A.D. dimensions. The regions in the P.A.D. emotional space represent human personality traits. Based on these regions, different personality traits of a person can be computed. As in BF personality traits validation, multiple scenarios must be generated to validate the performance of the temperament framework.

Furthermore, the robot should use the personality information and adapt its behaviour. There are many ways in which the robot can express its behaviour such as, gestures, postures, expressions, gaze and speech. Using personality theories from psychology, the robot should adapt appropriately. As mentioned earlier, robot can use the similarity-attraction principle and behave with the same personality type, or it can use the complementary attraction theory to adapt its behaviour and complement the personality of the interaction partner.

Next chapter (Chapter 5) discusses the visual perceptual system of the robot, while Chapter 6 presents the personality traits assessment system developed in this thesis.

5. Visual Perceptual System

Perception is the organization, identification, and interpretation of sensory information in order to represent and understand the presented information. Perceptual skills of a human are the abilities that enable a person to perceive, recognize and understand objects, humans, or a concept in the surrounding environment, and uses these skills to solve several tasks. In the scope of this thesis, the perceptive skills of a social robot are the abilities that enable a robot to perceive, recognize and understand human nonverbal cues in the surrounding environment in order to flourish human-robot interaction.

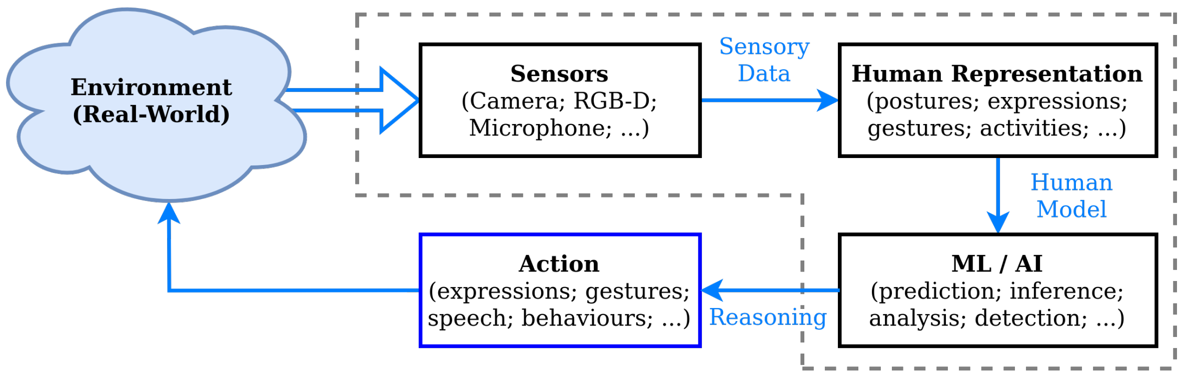


Figure 5.1: Key components of an interactive social robot's perceptual system: sensory data processing (visual and verbal perception); human representation based on a descriptor that contains all human information; algorithms to analyse, recognize, detect or assess different human features using AI/ML methods; and exhibiting behaviours for human-robot interaction

With respect to the field of human-robot interaction, fundamentally, the key components of a perceptual system of a robot are sensory data processing, data representation, and machine learning-based algorithms for prediction or inference. Using those components, a social robot understands its counterparts and executes appropriate behaviours and exhibits different reactions, as illustrated in Figure 5.1. Robotic cognition is highly significant for a social robot to understand humans and behave socially in real-world as expected

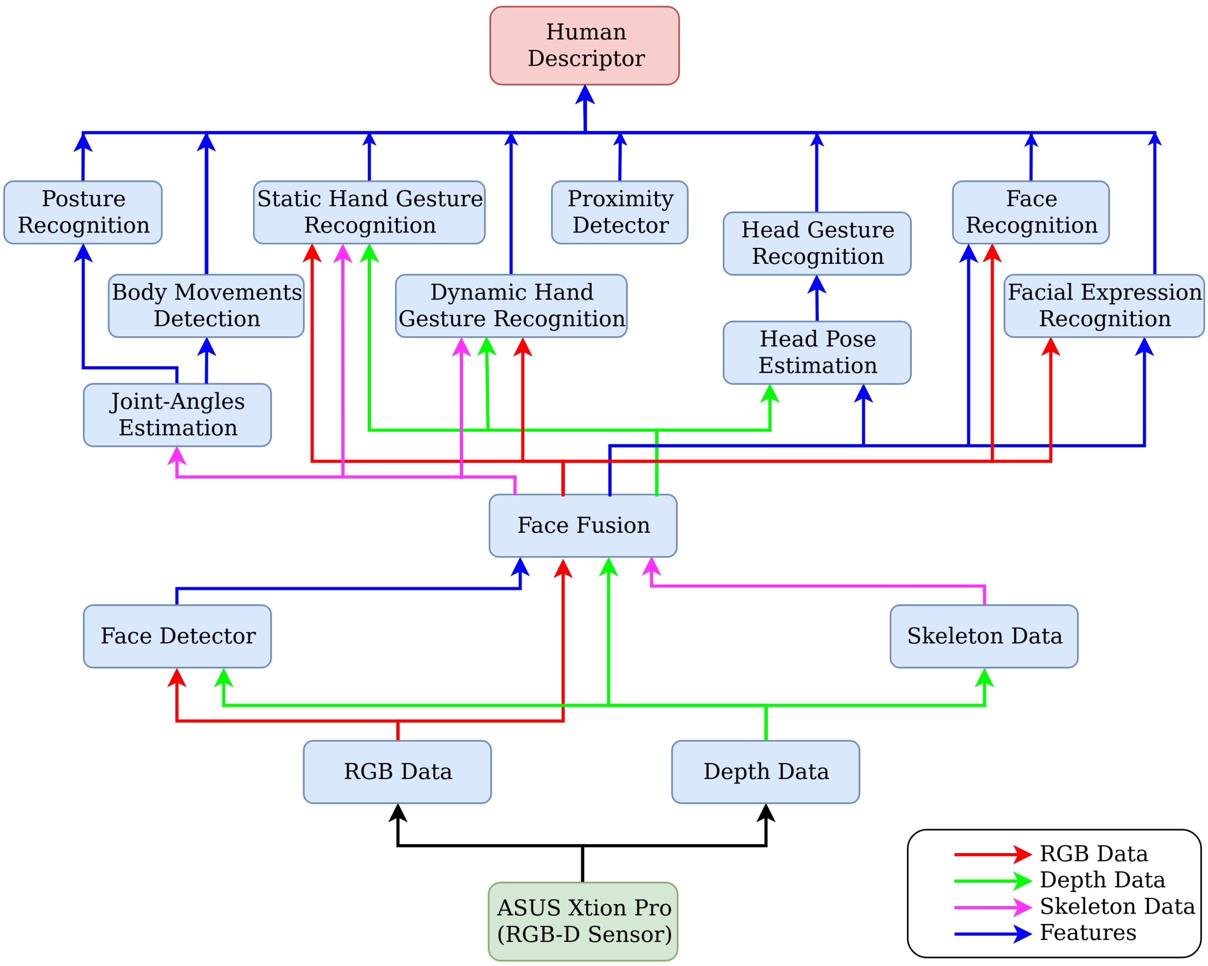


Figure 5.2: The visual perception system of an interactive social robot, ROBIN.

by society. However, the development of perceptual skills is a real challenge. Although some systems have been presented in the literature with such perceptual abilities, they perform poorly in the real-world environment. These systems mostly consider recognizing extremely naïve cues such as detecting faces or gestures. However, other human nonverbal features, such as proximity, speech duration, head gestures, postures and body movements also contribute a huge amount of information towards human behaviour understanding during social interactions.

This chapter focuses on the perceptual abilities that a robot must have in order to assess human personality traits. As discussed in previous chapters, assessing human personality traits helps the robot to maintain natural interaction with the interlocutor and also helps in self-adaptation of emotions and behaviours. This chapter introduces the visual perceptual system, in which the information from the sensor is processed to recognize several nonverbal cues. The sensor used in this work is an RGB-D sensor that presents colour (RGB) and depth image streams. The function of the perceptual system is to recognize and detect several *percepts*, such as faces, postures, expressions, static and dynamic hand gestures, head poses, and so on in real-time. Figure 5.2 shows the information processing of the visual perceptual system of the robot, ROBIN. All the features recognized in this system are fused together to form a single human descriptor which is used for assessment of human personality traits.

Perception of nonverbal features includes analysing and interpreting human behaviours. Most of the human behaviours are based on verbal and nonverbal cues. According to several studies from psychology, nonverbal cues convey a lot more information than verbal cues. Analysis of nonverbal cues is an important skill for a social robot to have, and it is also a challenging task for a robot to gather these skills.

5.1 Human Skeleton and Joints Information

Instead of using a monocular camera, ASUS Xtion Pro, an RGB-D sensor, is employed in order to exploit depth data along with colour images. The advantage of using such devices with depth sensor lies in the segmentation of the human skeleton using OpenNI and NiTE Library (see Appendix C for more details). Segmenting humans based on silhouette and edges may work in a constraint scenario, but it behaves poorly when applied in a dynamic environment. In contrast, a human can be detected and tracked efficiently using a depth sensor in a constantly changing scenario with a lot of different daily life objects involved. This sensor can work efficiently in the range of 0.5 to 4 meter.

Fifteen different skeletal joint positions of human can be extracted in real-time. These joint positions are quite accurate if the human is clearly visible. However, the extracted joint positions can be erroneous in some situations. This can happen in the following cases.

- The person limb(s) is/are occluded.
- Person's hands are too close to the body. The tracker is unable to distinguish between the torso and the limbs, and considers them as one whole object and erroneously fits the skeleton over it.
- Full length of human is not visible. It happens when the person comes too close to the robot.
- If the person is wearing big clothes, such as a jacket or a coat, that hides the limbs and the body silhouette, the skeleton tracker can also fail to accurately fit the skeleton.

The skeleton tracker can also track multiple humans and extract their joint positions in real-time. In order to extract joint positions reliably, the whole human body should be clearly visible to the sensor with no complete occlusions of body parts. The disadvantage in using joint positions is the dependence on correct detection of a human skeleton. Due to partial occlusions of limbs, the module can report ambiguous skeletal information which can effect the further percepts which are based on it. Figure 5.3 shows the tracked human with his skeleton visible. The outcome of this module is a percept, $P_{skeleton}$, which consists of 15 joint positions of human body as shown in Equation 5.1.

$$P_{skeleton} = ((x_1, y_1, z_1), \dots, (x_{15}, y_{15}, z_{15})) \qquad x_i, y_i, z_i \in \mathbb{R} \tag{5.1}$$

Figure 5.3: Tracked human with his skeleton information.

5.2 Face Detection

It is the fundamental and essential step in the analysis of face. The robustness of the face detector is the key aspect of our approach. There are several approaches for face detection [Hjelmas 01] and almost all of them detect a near-frontal or near-profile face. The viola-jones face detector is perhaps the robust real-time face detector and employed by many in the study of the face. This face detector consists of a cascade of classifiers trained by AdaBoost. Haar-like features are used on the integral image which can be computed very fast at any location of the integral image [Viola 04]. The performance of the face detector depends upon the number of training images. However, the Viola-Jones face detector does not accommodate high rotations of the faces.

The current work makes use of Haar cascade classifier to search for possible faces in 2D colour images. The faces detected by the Haar cascade classifier do not always represent a real face. It is due to the used 2D features that can be found in many variations of intensities (colours) analogous to a human face. It can not differentiate between a face in the picture and the real face of a person, which indicates a problem at the level of human-robot interaction. The corresponding depth information acquired by an RGB-D sensor is used to reduce the false positives that Haar cascade classifier produces [Saleh 13]. The output of the face detection module is a percept, P_{face}, which consists of 3D positions of the face.

$$P_{face} = (x, y, z) \qquad x, y, z \in \mathbb{R} \tag{5.2}$$

(a) 2D Face detection **(b)** 3D Face detection

Figure 5.4: Selection of real faces using face detection 2D and 3D. (Picture taken from [Saleh 13] p.133)

5.3 Posture Recognition

Human posture recognition is an active research topic in the field of human-robot interaction. In addition to being used in the context of humanoid robotics, recognition of human postures has many applications in human assistive systems and the automobile industry. The basic objective is to enable humanoid robots to work side by side with humans during daily life. In order to realize this goal, robotic systems must have the capability to differentiate human(s) from the cluttered environments. In addition to detecting humans, these systems should also analyze their posture, actions, emotions, motives and overall behaviour. It, in turn, helps the robots to be more intelligent and resourceful when interacting with humans. Human behaviour can be analyzed using human's nonverbal communication. According to [Noller 06], two-thirds of our communication consists of non-verbal communication, and only one-third of our communication consists of verbal content. Nonverbal communication consists of facial expressions and bodily cues. Human posture represents an important part of nonverbal communication.

Human posture and body movement play a significant role in the perception of the interaction partner. Humans use different hand gestures and body postures to express their internal emotional state in different situations. In humans, postures provide vital information through nonverbal cues. Psychological studies have also demonstrated the effects of body posture on emotions. This research can be traced back to Charles Darwin's studies of emotion and movement in humans and animals [Bruyn 03]. An extensive study and research have been conducted in the 1970s on the significance of body language in which the main area of focus was leg-crossing, defensive posture and arm-crossing, suggesting that these nonverbal behaviours depict feelings and attitude. Posture can also rely on the situation, i.e., people change their postures depending on the situation. Currently, many studies have shown that certain patterns of body movements are indicative of specific emotions [Dael 12] [Montepare 99]. Researchers have studied sign language and found that even non-sign language users can determine emotions from only hand and body movements [Rossberg-Gempton 93]. For example, anger is characterized by forwarding body movement [Oosterwijk 09].

Posture recognition plays an important role in expressing human emotions. Many researchers believe that all the variations of postures are due to the change in emotions and play a significant role in human evolution. Human emotions are always challenging to understand, and many factors influence human emotions. The art of recognizing human emotions had gained its importance long back and is, currently, studied actively [Al-Shawaf 16]. Some behavioural cues can be easily recognized from postures. For example, a person scratching his head during interaction shows thinking behaviour. Similarly, crossed arms posture shows that the interlocutor is reserved and is trying to block himself from opening to other people.

However, the challenge is to recognize complicated human postures in a cluttered environment in real-time, especially, those set of postures which are used in daily life in human-human interaction scenarios. For example, crossed arm, pointing with the left or right arm, casual or attentive standing postures, relaxing posture, thinking or shrug posture, etc. On the contrary, every region or culture has its different postures which, sometimes, are opposite in meaning in some another culture. One of the major challenges in recognizing human postures is diversity in people performing postures. People from a different culture are expressing the same posture in different ways as compared to others. In addition, postures are also dependent on the height and human physique variations which make them more challenging to recognize. Moreover, sitting postures appear different from standing postures and need separate classifier for the posture recognition task.

Numerous ways have been reported in the literature to recognize human postures. Some of these methods use wearable sensors to extract the psychological parameters like electroencephalography (EEG) data, skin temperature, accelerometer readings, etc. However, these methods require special sensors to wear all the time and sometimes require training on how to use them. In contrast, approaches using visual information from the visual sensors are more natural means of recognition of human posture. In the following section, we discuss the related work in the field of posture recognition.

5.3.1 Related Work

Research on posture recognition using skeleton data began in the 1990s and is still being carried on. Generally, posture recognition approaches can be separated into two broad categories: (a) wearable sensors based posture recognition and (b) posture recognition using vision-based sensors. Wearable sensors include gloves and other commercially available products that are used to extract different statistical and geometrical information of the limbs or body when worn. Few of these devices, namely Sensewear, ActiGraph and ActivPal, have been used by Wang et al. [Wang 16]. They address challenges, such as data imbalancement, instant recognition and sensor deployment in order to achieve an overall accuracy of 91% for sitting, standing and walking postures. Similar approaches using wearable sensors have been reported with higher accuracy. However, these require sensors to be worn. Latter approaches use vision sensors for the recognition of human postures. The advantage of this approach is twofold: first, these approaches are non-invasive; and secondly, they are also cost-efficient. Humans can perform their gestures and postures in front of a camera sensor without any other device attached to their bodies for posture recognition tasks.

Posture recognition via vision sensors can be further divided into two categories, namely camera-based posture recognition and RGB-D sensor-based posture recognition. Numerous works have been reported in the literature that uses a monocular camera to estimate human pose and human action. The most general approach is to extract features from images based on the structure of the human body, e.g., skin colour or face position [Lee 06b]. However, this approach imposes restrictions on features such as clothes and orientation. There are other methods to extract silhouettes and edges as features from the image [Agarwal 04] [Malik 02]. However, they rely on the stable extraction of the silhouettes and edges. Moreover, they perform poorly in self-occlusion.

In order to address these shortcomings, researchers use depth sensors to extract human joint positions. S. Nirjon et al. [Nirjon 14] describes a system, called Kintense, which is a real-time system with high accuracy to detect aggressive human actions, e.g., hitting and pushing that are relevant for games. The system has been trained using supervised and unsupervised machine learning techniques. The sensors calculate the distance between the body and the cameras, skeleton joints and speed at which an action is performed. Deep learning and neural networks are used to eliminate false positives and to identify actions that are not labelled. Real-time testing has been performed by deploying the system in more than one multiple-person household, which illustrates the sensitivity of the system towards unknown and unseen actions. The real-time system proves that the accuracy of the system is more than 90% [Nirjon 14].

Using RGB-D sensor, Zhang et al. [Zhang 14] extract joint positions of a human with the help of Microsoft Kinect. In order to make it independent of human size, each joint position is normalized using its neighbouring joint to make a feature. This feature vector which consists of all normalized joints, is then classified using SVM. A total of 22 postures are recognized with 3 different classifiers. The drawback of this approach lies in the normalization of joint positions. Although authors claim that the system is invariant to human size, it would not be invariant to human height or size of the limbs completely as normalization only adjusts joint values with its neighbouring joint.

Another similar work has been conducted by Ivan Lilloa et al. [Lillo 17] to recognize human activities using body poses estimated from RGB-D data. The system modules are classified into three different levels which include geometry and motion descriptors at the lowest level, sparse compositions of these body movements at the intermediate level, spatial and time-stamped compositions used to represent human actions involving multiple activities at the highest level. The work is related to the dictionary learning method, and their framework focuses on vector quantization using k-means to cluster low-level keypoint descriptors for dictionary learning [Boureau 10]. The model developed uses an alternative quantization method, discriminative dictionaries, or different pooling schemes [Niebles 10]. Sparse coding methods have also been used for alternative quantization methods. These methods have mostly focused on non-hierarchical cases where mid-level dictionaries and top-level classifiers are trained independently [Boureau 10]. Niebles et al. [Niebles 10] extend this model to the case of action recognition. In contrast to the former approach, the model is limited to binary classification problems and reports good accuracy only in a constraint scenario.

In previous related work, the required data is captured either from images or videos and the processing is done to create the feature vector. Feature vector represents the data in a form such that the system can be trained. Many classification techniques have been used in the classification of the training dataset, such as SVMs, neural networks and deep learning techniques. After the classification, the system can be tested offline using an existing database or online testing in a real-time scenario. Most of these approaches are used only to recognize standing postures or actions. Additionally, these approaches are not robust to real-time recognition of human postures with more than 10 classes. In this thesis, we have proposed an approach that is robust to real-time recognition of postures and can differentiate between standing and sitting postures. Moreover, it can recognize 19 postures used in daily life routine. The detailed analysis of the implemented approach is discussed in the following sub-sections.

5.3.2 Methodology

Visual perception in complex and dynamical scenes with the cluttered background is a challenging task which humans can solve remarkably well. However, it performs poorly in this kind of challenging scenarios for a robot perception system. One of the reasons for this significant difference in performance is the use of context or contextual information by humans. Furthermore, the robot has to perform its computations as fast as possible due to the notion of real-time. As a result, most of the time, the robot perception system is hampered with low-resolution images. There is a need to develop such a recognition system which can recognize postures in complex environments and work efficiently.

This work presents an approach that uses depth data to extract skeletal human joint positions. These joint positions are then converted into meaningful angles for feature vector generation task. The resultant feature vector is unique for each posture and is invariant to height, body shape, illumination, proximity and appearance of a human. The working schematics of the proposed approach is presented in the Figure 5.5. The approach reports high accuracy for both sitting and standing postures. The system can recognize overall 19 gestures real-time when classified by using multi-class SVM. Figure 5.8 shows different postures for standing that are recognized by the system. The system also recognizes similar postures for sitting. Each module of the approach is described in the following sub-sections.

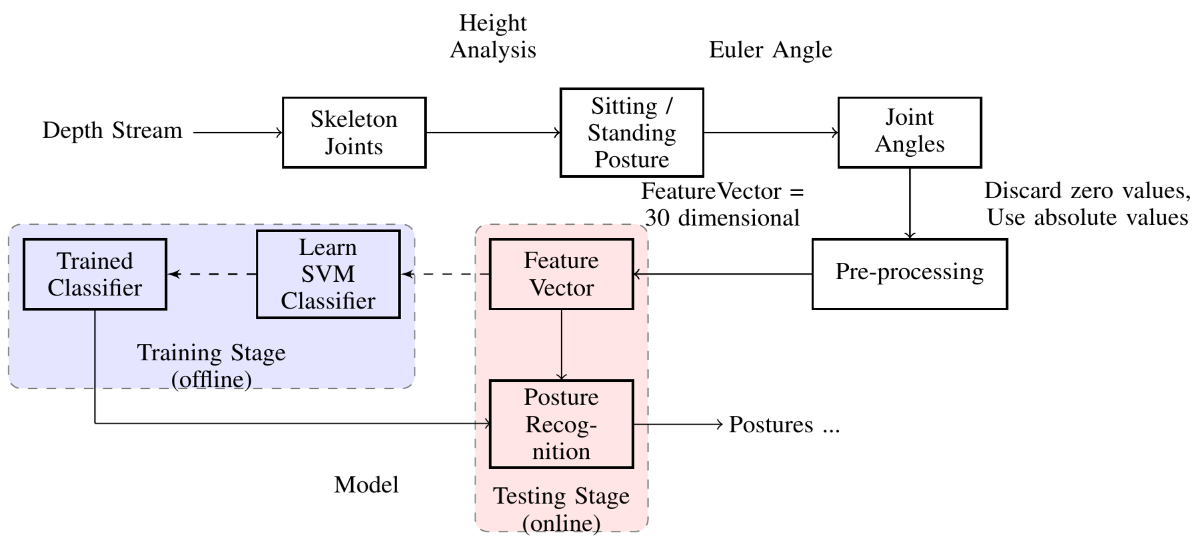


Figure 5.5: Working schematics of the approach. Using depth stream, skeleton joints are extracted. Based on the height analysis, the system classifies the subject either standing or sitting. After which joint angles are computed from joint positions to construct a 30-dimensional feature vector for classification.

Detection of Posture Type

Before recognizing postures, the important step is to detect whether a human is standing or sitting. The simplest way is to analyse the height of human with respect to its z distance from the sensor. Empirical studies have shown that the relation between these two entities is linear. For example, if a human is near to the sensor, he/she appears taller, and similarly, if a human is away from the sensor, he/she appears short.

To make it height and scale invariant, the system uses the depth data (z distance) to normalize the height of the person. If the human head joint has the value more than the set threshold value, δ, the system classifies it as a standing posture. If he/she has the head joint position value less than the set threshold, the system classifies it as sitting.

The outcome of this module is a percept, $P_{posture\_type}$, which has the information about the type of posture, as illustrated in Equation 5.3. Algorithm 5.1 shows the process of detection of posture type.

$$P_{posture\_type} = (pt) \qquad pt \in \{standing, sitting\} \tag{5.3}$$

Algorithm 5.1: Detection of posture type using height analysis.

```
Standing← false;
Sitting← false;
if Human.Body.Head_Joint().Z ≥ δ then
    Standing← true;
else
    Sitting← true;
end
```

Joint Angles Estimation

The major disadvantage in using joint positions for feature extraction task is that they are variant to positions, height and limbs variations. This type of features might report better results when the position and height of the human are fixed. However, these features behave poorly when dealing with varied height or dynamic humans. In order to solve this problem, researchers calculate the difference of each joint from torso joint to construct a feature vector. This ensures that the features are independent of human positions and movements. Despite being better than previous method, it is still dependent on the height of the person.

In order to address this shortcoming, this thesis proposes a unique method to extract features. Instead of using joint position for feature extraction task, these joints positions can be converted into angles (roll, pitch, yaw) between every two joints for construction of feature vector. The benefit of using angles is that they are not dependent on the position or height or human physique. Instead, they compute directions between each joint. The direction between each joint would be similar for a short person and a tall person if they

Algorithm 5.2: Angle estimation between the difference vectors of two joints.

Consider 'a' and 'b' are two joints;
while *Human.Body.Exist()* **do**
- $angle_{xy} = \tan^{-1} \frac{(a_y - b_y)}{(a_x - b_x)}$;
- $angle_{yz} = \tan^{-1} \frac{(a_z - b_z)}{(a_y - b_y)}$;
- $angle_{zx} = \tan^{-1} \frac{(a_x - b_x)}{(a_z - b_z)}$;
- `/* Converting angles from radian to degrees */`
- $angle_x = angle_{yz} \times 180/\pi$;
- $angle_y = angle_{zx} \times 180/\pi$;
- $angle_z = angle_{xy} \times 180/\pi$;

end

are expressing the same posture. Euler angles are used to convert the joint positions into angles. Algorithm 5.2 shows the calculation of angle between the difference vectors of two joints a and b. The output of this sub-module is another percept, P_{joint_angles}, which has the information of all the 15 joint angles, as shown in Equation 5.4.

$$P_{joint\_angles} = ((x_1, y_1, z_1), \ldots, (x_{15}, y_{15}, z_{15})) \qquad x_i, y_i, z_i \in \{0 - 360°\} \tag{5.4}$$

Pre-processing and Feature Extraction

In total, 15 joint angles can be calculated for each posture. However, it has been observed during experiments that certain joints do not contribute to the posture. Joint angles between knee and foot, or hip and knee do not add useful information for the posture recognition task. The reason lies in the postures recognized in this work, which are not affected by joint angles of the lower body. During this pre-processing stage, the number of angles recorded is reduced to 10.

After empirical studies, it has been found out that when part of the limb or body is occluded, skeleton tracker reports false joint positions, $(0, 0, 0)$, which creates a problem for the classification task. Training on erroneous data leads to poor classification and the accuracy of the system goes down. Therefore, if the joint positions with values $(0, 0, 0)$ appear in 10 consecutive frames, these instances are discarded. 10 joint angles are then used to construct a feature vector. Since every joint angle is 3-dimensional, the feature vector for a single depth observation becomes 30 dimensional.

Classification

Classification is an important step in any recognition task. The major task of the classification stage is to differentiate each class or category accurately based on the knowledge gained during the training stage. Numerous classification algorithms have been presented in machine learning, e.g., neural networks, decision trees, random forests, convolution neural network, etc. This module uses SVM, a supervised learning algorithm, for the classification task. An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall [Tong 01]. The benefit of SVM lies in the regularization parameter, which if set accurately, avoids over-fitting. Moreover, it uses the kernel trick, i.e., it can build an expert knowledge about the problem by engineering the kernel. SVM generalizes on the high dimensional feature set quite well given that the database is also huge.

This work uses a multi-class SVM classification. The reason of selecting SVMs is that the features coming out after feature extraction are quite distinct and linearly distinguishable (see Figure 5.6 and Figure 5.7). Moreover, SVM classification using linear kernel is highly efficient as compared to some other classification algorithms such as convolutional neural networks. In order to make the task computationally intensive, we have selected to use SVMs.

More than 2100 instances are used during the training stage for each posture, and 40000 instances for the whole training data are used for 19 classes. Figure 5.8 shows different

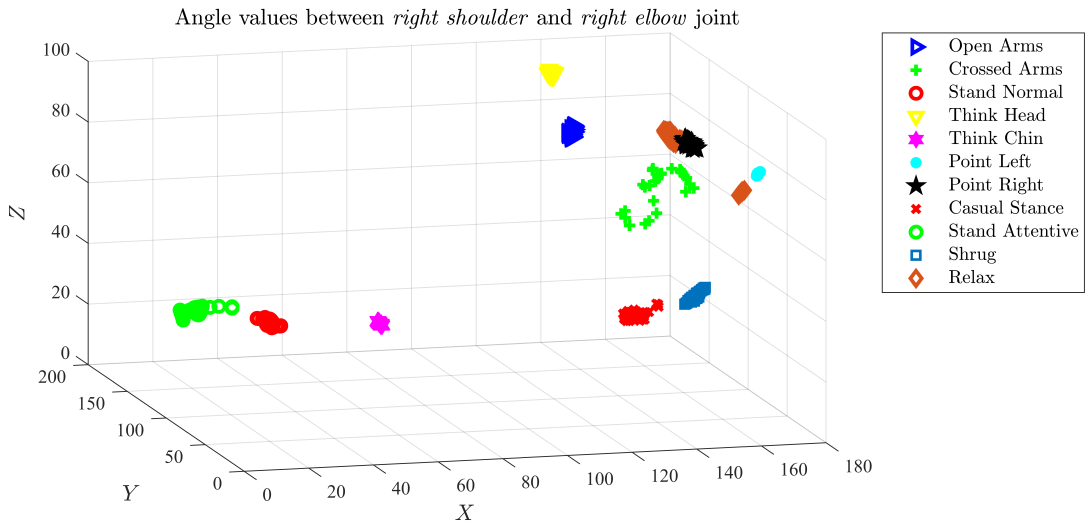


Figure 5.6: Angle values between right shoulder and right elbow from training data in 3D plane. Each class is depicted with different colour.

postures for standing that are recognized by the system. The system also recognizes similar postures for sitting. 10 different subjects, from different ethnicity (Indian, Pakistani, German, Italian and Turkish), featured in the training dataset. In spite of subjects from different ethnic backgrounds, it has been found out that all the subjects have performed the postures in the similar way. One of the possible reason is the type of postures selected in this study which are quite standard. A linear kernel with regularization parameter $C = 0.4586$ is used during SVM training. Figure 5.6 and Figure 5.7 show 3D graphical plots of joint angles between *right_shoulder-right_elbow* and *right_elbow-right_hand*, respectively. It can be seen that the classes are easily distinguishable based on the angle between the two joints. With the contribution of other joints angles between joints, the problem is easily classified by the SVM linear kernel.

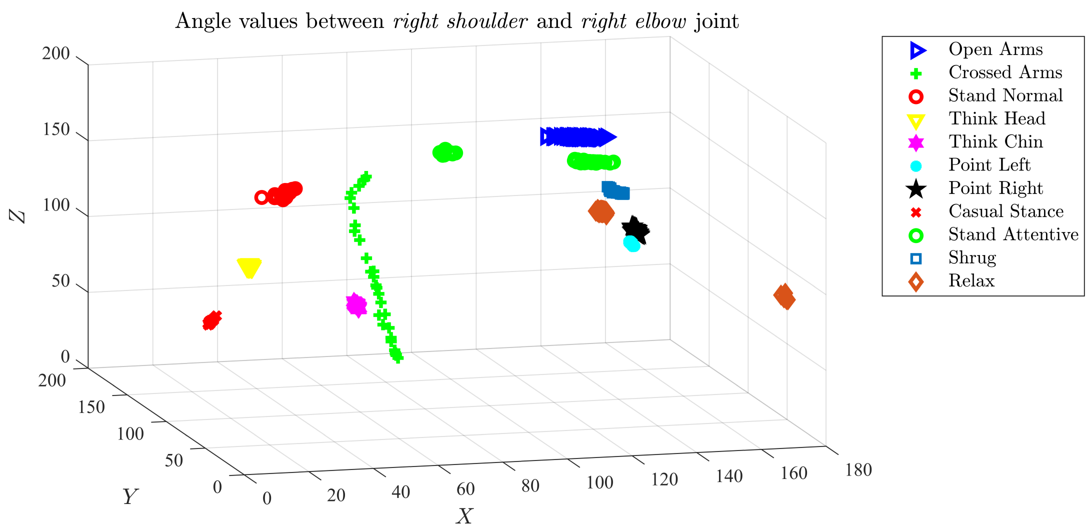
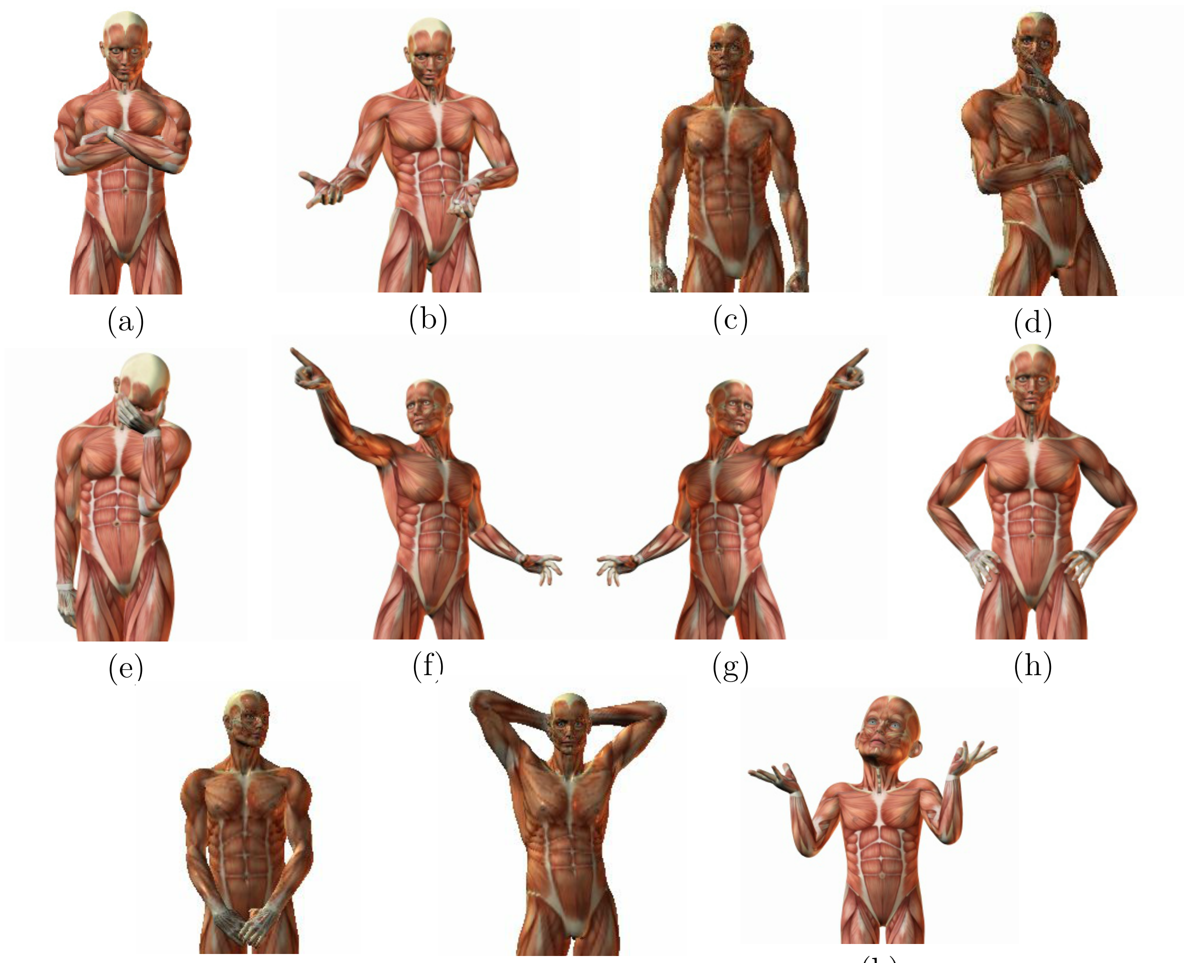


Figure 5.7: Angle values between right elbow and right hand from training data in 3D plane. Each class is depicted with different colour.



(a) (b) (c) (d) (e) (f) (g) (h) (i) (j) (k)

Figure 5.8: Standing postures (a) Crossed Arms (b) Open Arms (c) Stand Normal (d) Think (Hand on chin) (e) Think (Hand on Head) (f) Point Right (g) Point Left (h) Casual Stance (i) Attentive (j) Relax (k) Shrug. (pictures are used from `www.posemaniacs.com`)

5.3.3 Experimentation and Evaluation

The goal of the posture module is to recognize human postures in real-time robustly in order to realize human-robot interaction. Postures are categorized mainly as sitting and standing. Overall, 11 postures are recognized for standing, and 8 postures have been recognized for sitting. Different postures recognized in this thesis are crossed arms, open arms, think (hand on the head), think (hand on chin), pointing (with left hand), pointing (with right hand), standing/sitting normal, shrug, relax, casual posture and attentive posture, as shown in Figure 5.8.

These postures are selected based on their significance for the assessment of human personality traits. According to [Kuhnke 12], fist slamming, sharp finger-pointing and stomping feet on the ground are all positively correlated with self-centred personality trait. A body's forward and erect posture with feet wide apart reveals aggression and is positively correlated with self-centred trait. Moreover, open and confident posture is the norm for individuals in high-status positions such as leaders and also positively correlated with self-centred trait. Furthermore, introverts tend to block others by crossing their arms on their chest in a way to build a barrier. They also duck their heads and look away [Kuhnke 12]. A total of 10 subjects have featured in the training stage. For each class and each subject, at least 300 instances are collected with a little bit of movement and varied styles.

Experiments

There are generally two ways to conduct experiments. Experimentation or testing of the system can either be done on the testing dataset or testing can be done real-time directly on the robot, ROBIN. We have conducted both these experiments in this work to evaluate the system. 25% of the dataset have been separated from the training data before training. This dataset serves as a test dataset to evaluate the system. Since the recorded dataset has no false skeleton tracking, the system reports 99.4% recognition rate. It shows the potential of the approach when the provided dataset is accurate.

Table 5.1: Standing Postures and their Recognition Rates

	Postures	**Recognition Rate (%)**
1	Crossed Arms	100
2	Open Arms	100
3	Standing Normal	97.33
4	Think (Hand on the Head)	98.67
5	Think (Hand on the Chin)	95.1
6	Point with Left Hand	100
7	Point with Right Hand	100
8	Casual Stance	100
9	Standing Attentive	90.27
10	Shrug	100
11	Relax (Hands behind the neck)	100
	Average	**98.3**

For the second experiment, social robot, ROBIN, is used to recognize postures in real-time. Once ROBIN recognizes the posture, it indicates by saying the name of the posture. In order to avoid any bias, new subjects have been used to express postures in front of ROBIN. Subjects have been instructed in the start about the postures which ROBIN can recognize. However, the knowledge about performing each posture has not been shared with them in order to evaluate the system potential to generalize varied postures. Every subject performs each posture at least 30 times. Table 5.1 and Table 5.2 show the recognition rates of standing and sitting postures respectively.

Performance Evaluation

As shown in Table 5.1 and Table 5.2, ROBIN is able to recognize human postures with an average accuracy of 98%. For standing postures, the recognition rate for each class is above 95% except for the ‘standing attentive’ posture. Attentive posture reports low accuracy as compared to others because the hands are too close to the body, and therefore, the algorithm considers hands as part of the body for skeletal joints detection. Thinking

Table 5.2: Sitting Postures and their Recognition Rate

	Sitting Postures	**Recognition Rate (%)**
12	Sitting Normal	98.33
13	Crossed Arms	95.28
14	Think (Hand on the Head)	96
15	Think (Hand on the Chin)	94.5
16	Point with Left Hand	100
17	Point with Right Hand	100
18	Shrug	100
19	Relax (Hands behind the neck)	100
	Average	**98.01**

postures are sometimes confused between each other, but nevertheless, show recognition rates above 95%.

For sitting postures, it has been found out that when the person is sitting, the skeleton of the whole body is not visible. In order to address this issue, ROBIN uses torso pitch angle to tilt its body in the front. In this way, the whole skeleton of human is visible. Due to a sitting posture, sometimes human skeleton tracker does not work accurately to localize limbs and positions. Therefore, some of the postures show relatively less recognition rate than standing postures. Nevertheless, ROBIN can recognize human postures accurately in real-time with an accuracy of more than 98%. Since the system uses only depth data, issues regarding lighting condition, image resolution, texture variations are avoided. It enhances the accuracy considerably as compared to approaches using colour image to recognize human postures. Figure 5.9 shows an experimental environment where the subject is interacting with ROBIN using postures. The outcome of the posture module is a percept, $P_{posture}$, which represents postures, posture ID and posture type, as illustrate in Equation 5.5.

$$
\begin{aligned}
P_{posture} = (id, p, pt) \qquad & id \in \mathbb{N} \; ; \\
& p \in \{crossed\_arms, open\_arms, pointing, thinking\_head, relaxing, \\
& \qquad thinking\_chin, shrug, standing\_normal, aggressive\_stance\} \\
& pt \in \{standing, sitting\}
\end{aligned}
\tag{5.5}
$$

5.4 Proximity and Body Movements Detection

Proxemics is the study of using spatial distances that individuals maintain in interpersonal communication and how these distances are affected by environmental and cultural factors. Two important aspects are of interest in studying proxemics: personal space and interpersonal distance. Personal space is a portable area immediately surrounding the body, within which human moves around. It can grow or shrink depending on human personality and the situation in which human interacts. Human controls who is permitted to get inside this area and who is not.

Interpersonal distance is the distance that the interaction partner maintains during their conversations. This distance conveys important messages to both of the interactants and

Figure 5.9: Subject is interacting with ROBIN using postures.

other people. If the distance between two interaction partners exceeds or is less than the limits which are predefined by environmental or cultural factors, then negative attitudes are elicited. Mehrabian [Mehrabian 69] stated that the distance between an individual and his addressee is a decreasing linear function of the degree of liking of the addressee. Moreover, individuals high on extroversion prefer to sit or stand close to the conversation partner [Knapp 13] [Hargie 16]. According to Hall [Hall 63], interpersonal distances of a person can be categorized into four zones, namely intimate space, personal space, social space, and public space.

Proximity information can be extracted by using the depth data of tracked human. A change in stance near or farther from the robot shows comfortability of a person. If a person is an extrovert, he is more likely to move closer to interlocutor and vice versa. The distance between the robot and the person is a very crucial parameter to recognize personality trait. The robot is capable of measuring the distance between its camera and the position of the person. When there is a positive difference between the two positions, the person comes nearer; otherwise, the person goes farther. However, if there is no difference, this feature has no impact on personality assessment. In this thesis, the difference of proximity, $\triangle P$, in the z direction has been calculated by subtracting the final proximity from the initial proximity in the z direction. Three types of stances, namely forward stance, backward stance and neutral stance, have been considered in this work. If the difference, $\triangle P$, in z direction is $> 10cm$, the forward stance is activated. On the contrary, if the difference has a value less than $-10cm$, it is termed as 'backward stance'. If these two conditions are not satisfied, the stance is termed as 'neutral stance' or 'no stance'. Based on the type of stance displayed by the communicating partner during a conversation, the proximity estimation value is obtained. If the person is found out to be showing a high rate of forwarding stance, the value for this feature is flagged. Algorithm 5.3 shows

Algorithm 5.3: Proximity estimation of a person with respect to the robot.

```
distance_t denotes Human.Face.Center().Z value at time t;
distance_{t-2} denotes Human.Face.Center().Z at time t − 2 (2 seconds);
/* where Human.Face.Center().Z is the distance between the robot and the
   person's face.                                                     */
△P ← change in proximity;
while Human.Face.Exist() do
    △P = distance_{t-2} − distance_t;
    if diff ≥ 10cm then
        proximity = forward_stance;
    else if diff ≤ −10cm then
        proximity = backward_stance;
    else
        proximity = neutral_stance;
    end
end
```

the process of calculating human stance with respect to the robot.

Similarly, body movements during an interaction is an important nonverbal cue which sheds light on the spirit of a person. These movements are related to a person's limbs and body in general. Excessive hand movements, along with other body cues, is a sign of confidence [Oberzaucher 08]. According to Nass et al. [Nass 01], people, when aroused, show frequent body movements. Extroversion is related to more frequent and more rapid body movements [Oberzaucher 08].

Algorithm 5.4: Activity estimation of a person.

```
Angles_t denotes P_joint_angles vector at time t;
Angles_{t-1} denotes P_joint_angles at time t − 1 (1 seconds);
/* where P_joint_angles are the angles between human joints.          */
while Human.Body.Exist() do
    diff = Angles_{t-1} − Angles_t;
    if diff ≥ δ then
        body_movements ← true;
    else
        body_movements ← false;
    end
end
```

To detect body movements, we compute skeletal joint angles, as discussed in section 5.3, of the upper body and analyse it over time. The change in angle values are recorded, and if it exceeds a threshold, δ, activity is flagged. Algorithm 5.4 shows the estimation of activity using joint angles. The result of these features (i.e., proximity and body movements) are two percepts that represents human stance and human activity, respectively. Equation 5.6 shows the percept, $P_{proximity}$, and Equation 5.7 shows the percept, $P_{body\_movements}$.

$$P_{proximity} = (z, proximity\_stance, \triangle P) \qquad \triangle P, z \in \mathbb{R} \; ; \quad proximity\_stance \in \{forward, neutral, backward\} \tag{5.6}$$

$$P_{body\_movements} = (activity) \qquad activity \in \{active, inactive\}] \tag{5.7}$$

5.5 Speech Duration

As mentioned in [Jensen 16], extroverts are more likely to be talkative. This feature can easily be correlated with the duration of the speech. A person that talks for long duration are known to be an extrovert and a person who talks less is termed as an introvert and a loner, who avoids people. Extroverts talk more and they produce more words and talk longer when they have the turn [Argyle 13]. Extroverts talk faster, louder, with shorter pauses and with a higher pitch [La-France 04] [Matsumoto 13] [Knapp 11]. We categorize the speech duration by analysing the time taken by the human in two classes, i.e., long duration and short duration based on a fixed threshold, δ. This threshold value is obtained from empirical experiments.

Due to the absence of speech recognition, the estimation of speech duration is quite challenging. The robot does not detect when the person has start or stopped talking. For this purpose, we can use two methods to send a signal to the robot that the person has stopped speaking. The first method is to specify a specific percept which if recognized by the system, communicates the robot that the person has finished speaking. In this case, we have used *raise hand* posture. When the person has stopped speaking, he/she can raise the hand to let the robot know he has finished speaking. In the second method, a second person, who is supervising the experiment, is responsible to let the robot know that the interaction partner has stopped speaking. A button has been added in the graphic user interface (GUI) of the robot which when pressed sends a signal to the speech duration module.

The speech duration starts as soon as the person starts responding and finishes when the person stops speaking. The final point of time is subtracted from the initial point of time to get the duration for a specific question asked by the robot. Since the context of the conversation is known, the normal duration is empirically set to 15 seconds. If the human takes more than this threshold, δ, to answer a predefined question, the speech duration is termed as long speech duration and vice versa. The result of this module is a percept, $P_{speech\_duration}$, which has the information about the duration of speech (i.e., long or short). Equation 5.8 shows the resultant percept, $P_{speech\_duration}$, while Algorithm 5.5 shows the process of estimating the duration of speech using 'raise_hand' percept.

$$P_{speech\_duration} = (duration) \qquad duration \in \{long, short\} \tag{5.8}$$

5.6 Facial Expressions Recognition

Facial expressions of humans play an important role in inter-human interaction. Usually, humans express more feelings through facial gestures than any other body movements.

Algorithm 5.5: Speech duration estimation of a person.

```
start_time ← Current.Time;
end_time = 0;
while Human.Body.Exist() do
    P_posture ← Human.Current.Posture;
    if P_posture = raise_hand then
        end_time ← Current.Time ;
        diff = end_time − start_time ;
        if diff ≥ δ then
            speech_duration ← long;
        else
            speech_duration ← short;
        end
    else
        continue;
    end
end
```

The facial expression also represents the internal emotional state of a person. It plays a vital role in explaining the intended meaning of the speech.

Because of the importance of face in emotion expression and perception, most of the vision-based affect recognition studies focus on facial expression analysis. Recognizing human emotions using facial expression is a common practice. In the field of HRI, facial expression recognition has got a lot of attention since the last decade. Many different approaches are used in recognizing facial expressions. As far as automatic facial affect recognition is concerned, most of the existing efforts studied the expressions of the six basic emotions due to their universal properties, their marked reference representation in our affective lives, and the availability of the relevant training and test material (e.g., [Kanade 00]). There are a few tentative efforts to detect non-basic affective states from deliberately displayed facial expressions, including fatigue [Gu 04], [Ji 06], and mental states, such as agreeing, concentrated, interested, thinking, confused, and frustrated (e.g., [Kaliouby 05], [Kapoor 07], [Kapoor 05]).

The current work uses the approach implemented by Aldarraji et al. [Al-Darraji 16a]. In their work, the authors use a convolutional neural network (CNN) to recognize related action units and their combinations on the face. Action unit is related to the contraction of one or more facial muscles. There is a total of 44 action units which can also occur in combination on a face. CNN consists of multiple layers that are connected with each other. These layers can be of different types such as a convolutional layer, pooling layer or fully connected layer.

Figure 5.10 shows the architecture of the convolutional neural network. This architecture is used to recognize different facial action units. As can be seen that the architecture consists of 6 layers apart from input layers. The architecture receives a 32×32 grey input image with a face and outputs the confidences of 6 basic facial expressions along with a neutral. It uses two convolutional layers, each with its own sub-sampling layer (max-pooling). The convolutional layer applies a set of learnable filters on the input image. Using more than one convolutional layer enables to extract features on different levels.

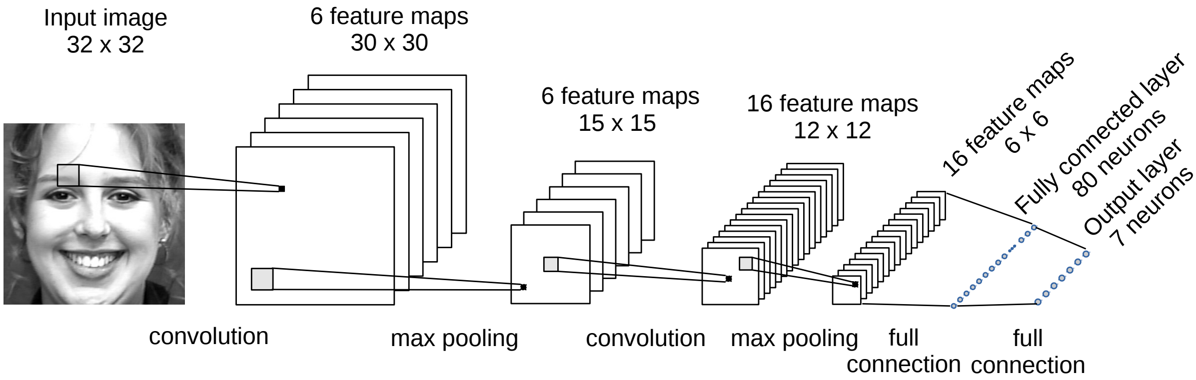


Figure 5.10: Architecture of deep neural network for facial expression recognition. It contains two convolutional layers each with its maximum sampling layer and one fully connected layers with 80 neurons that is connected to the output layer of 7 neurons (image from [Al-Darraji 16b] p.4).

Each convolutional layer extract higher-level features than the previous layers. For $m \times m$ filter w, the output unit x^l_{ij} of the convolutional layer l is calculated using equation 5.9.

$$x^l_{ij} = \sum_{a=0}^{m-1} \sum_{b=0}^{m-1} w_{ab} y^{l-1}_{(i+a)(j+b)} \tag{5.9}$$

The first convolution layer uses a kernel of size 3×3 which produces an image of 30×30 pixels and 6 channels. The max-pooling layer following the first layer uses a kernel of size 2×2, which reduces the image size to 15×15 pixels. The second convolution layer applies a kernel of size 4×4 that produces an image of 12×12 pixels and 16 channels. The last max-pooling layer uses a kernel of 2×2 which produces an image of 6×6 pixels. Finally, a fully connected layer of 80 neurons that receives the output of the two convolutions and in turn is connected to seven output neurons, which represents six emotions in addition to neutral. The activation function that is used in all layers is *tanh* function.

In the learning phase, mean squared error (MSE) is used as a loss function. The optimization function used in this work is the Levenberg-Marquardt gradient descent, which combines the advantages of the steepest descent method with the Gauss-Neuton method by adaptively varying the parameter updates between the two methods. The approach reports 90.85% for all the action units and 90.3% for 6 basic facial expressions, (see [Al-Darraji 16a] for more details). The resulting percept, $P_{facial\_expression}$, of this module is given in Equation 5.10 which has the information of 3D face coordinates and expression, e.

$$\begin{aligned} &P_{expression} = (x, y, z, e) \qquad x, y, z \in \mathbb{R} \ ; \\ &e \in \{neutral, angry, disgust, fear, happy, sad, surprise\} \end{aligned} \tag{5.10}$$

5.7 Head Gesture Recognition

Another nonverbal aspect, which has been implemented is human head movement or head gestures. Humans have the ability to interpret these movements quickly and effortlessly.

However, it is regarded as a difficult challenge in computer systems and robotics. Detecting human head movement requires estimating head pose over time. For example, head nodding is the deviation of the pitch angle of the head, whereas head shaking is the deviation of the yaw angle.

In order to build a reliable human-robot interaction system, a robust head pose estimation algorithm is needed. The current work utilizes head pose estimation used in [Saleh 13] and [Saleh 15]. Eight head gestures are recognized in this work. These gestures are *nodding*, *shaking*, *tilting*, *looking ahead*, *looking left*, *looking up*, and *looking down*. Head nodding and shaking are dynamic gestures, in which the head pose is changing over time, whereas the others are static gestures in which the head pose is nearly not changed.

Head nodding gesture can be detected, by the robot, as a sequence of poses where the *pitch* angle of the head exceeds the threshold θ_p in both directions. The speed of the nods depends on the number of nods and the time in which the nods occurred.

$$r_n = \frac{N}{t} \tag{5.11}$$

In Equation 5.11, N is the number of nods, t is the time in which the nods are occurred, and r_n is the nodding rate.

$$Nodding\ Speed = \begin{cases} slow & if\ r_n < \rho_n \\ fast & if\ r_n \geq \rho_n \end{cases} \tag{5.12}$$

In the above equation, ρ_n is the nodding speed threshold. Head shaking is detected in the same way but on the *yaw* angle instead of *pitch*.

Static gestures can be detected by calculating the duration of head poses in different directions. If the duration of a specific direction exceeds 80% of a specific period of time, then the corresponding gesture is regarded as *active*. Otherwise, the gesture is regarded as *inactive*. The percept, P_{head_pose}, has the information of human head poses, while the percept, $P_{head\_gesture}$, has the information of human head gestures, as shown in Equation 5.13 and Equation 5.14, respectively.

$$P_{head\_pose} = (x, y, z, \phi, \theta, \psi) \qquad x, y, z \in \mathbb{R}\ ;\ \phi, \theta, \psi \in [-90°, +90°] \tag{5.13}$$

$$\begin{aligned} &P_{head\_gesture} = (id, hg) \qquad id \in \mathbb{N}\ ; \\ &hg \in \{nodding, shaking, tilting, looking\} \end{aligned} \tag{5.14}$$

5.8 Hand Gesture Recognition

Hand gesture recognition has been a popular topic in the computer vision field. The topic has been studied numerous times because of its important applications in surveillance systems, elderly care, in the field of medicine (e.g., gait analysis, surgical navigation), in the field of sports, augmented reality, sign language for hearing impaired people and human behaviour analysis. Hand gestures are critical in face-to-face communication scenarios.

Especially during discussions, hand gestures become more animated. They emphasize points and convey the enthusiasm of the speaker. Hand gestures tell a lot about the internal state. For example, crossing arms during face-to-face communication show nervousness or lack of interest and clenched hands show aggressive stance of a person. They enable humans to express mood state (such as thumbs up, fist) or convey some basic cardinal information (such as one, two, and so on).

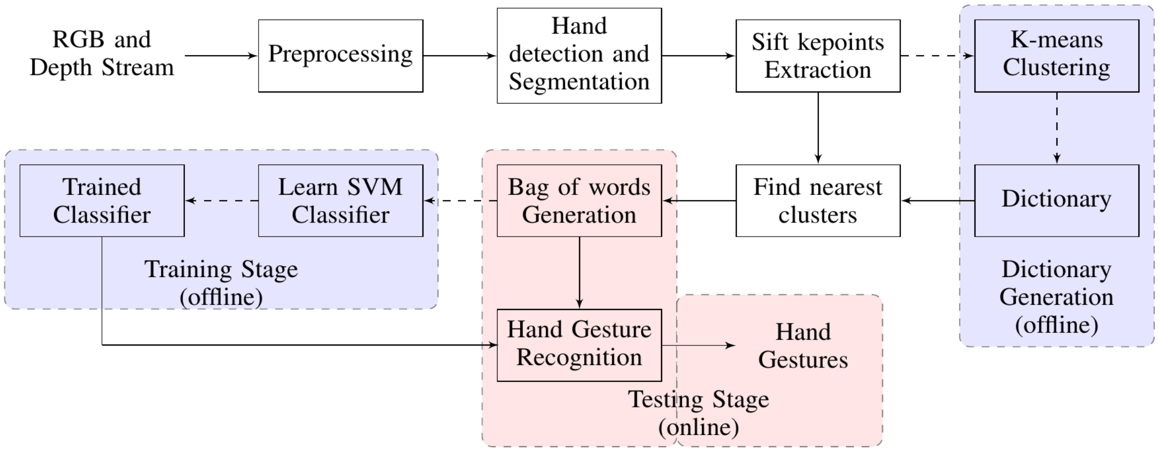


Figure 5.11: Working schematics of hand gesture approach. Using depth stream, the hand is localized and segmented, and then SIFT algorithm is used to extract keypoints. A dictionary is generated using k-means clustering; keypoints of every image is compared to find nearest clusters and bag-of-words are formed; which are fed into SVMs for training classifier in training stage; and for classification of different hand gestures in testing stage

Numerous hand gesture recognition systems have been reported in the literature. In general, we can categorize them in two different classes: (a) data gloves based systems and (b) vision-based systems. The former type of systems requires the use of glove sensor for storing hand and finger motion and then use this data to recognize the action. Huang et al. [Huang 11] used gloves to record the hand and fingers flex data and then use machine learning algorithms to classify 5-dimensional finger flex data. This type of systems might provide a 3-D representation of hand but wearing a heavy and expensive glove is not suitable for natural human interaction. On the other side, vision-based systems take the information of the hand itself as an input using a camera to collect hand movements for gesture recognition without the use of any wearable sensor. Vision-based approaches can be divided into two categories, i.e., 3-D hand model-based method and appearance-based methods. The 3-D hand model can provide ample information of hand that allows realizing wide class of hand gestures. Still, the main disadvantage lies in the extraction of features in case of ambiguous poses and unclear views and high computational complexity which makes the overall system unrealisable for real-time interaction.

Keeping this in mind, we present an appearance based approach that uses a depth sensor for localization and segmentation of hand and then, uses bag-of-words model by extracting SIFT keypoints and classifies them using support vector machines (SVMs). The block diagram of the approach is presented in Figure 5.11 (see [Zafar 16a] for more details).

Figure 5.12: Subject performing different gestures in front of the robot.

19 hand gestures have been recognized in this thesis which include open palm, victory, thumb up, thumb down, fist, inquire, point up, okay gesture and number gestures (i.e., one, two, three and so on). Figure 5.12 shows a subject performing different gestures. Although the mentioned gestures are important for human-robot interaction but due to their low significance toward personality trait assessment, hand gestures are not considered for personality trait task. Nevertheless, the hand gestures are used to interact with the robot and to answer. Table 5.3 shows the recognition rates for each hand gesture.

Table 5.3: Recognition Rates for Common Hand Gestures.

Hand Gesture	**Number of Images**	**Correctly Classified**	**Recog. Rate**
Okay gesture	150	115	76.6%
Pointing Up	150	149	99.3%
Thumbs Up	150	136	90.7%
Thumbs Down	150	123	82%
Inquiry gesture	150	141	94%
Fist	150	144	96%
Palm	150	150	100%
Victory	150	145	96.7%
Little finger gesture	150	137	91.3%
Average			**93%**

As shown in the Table 5.3, the average recognition rate in case of common hand gestures when performed in front of the robot is 93%. For Okay gesture, the recognition rate is low due to the reason that the fingers are pointing upwards, and sometimes it is confused with victory and inquire gesture. Gestures such as thumbs up and little finger are occasionally confused with pointing up gesture. In all three cases, one finger is pointing upwards, which results in false detection. An important thing to note is that all of the gestures are tested when the hands are dynamic. The outcome of the hand gesture module is a percept, $P_{hand\_gesture}$, which has the information of hand position and the gestures. Equation 5.15 shows the percept, $P_{hand\_gesture}$.

$$\begin{aligned} P_{hand\_gesture} = (x, y, z, g) \qquad & x, y, z \in \mathbb{R} \ ; \\ g \in \{ & thumb\_up, thumb\_down, okay, victory, inquire, point\_up, \\ & number_1, number_2, number_3, \ldots, number_{10}\} \end{aligned} \tag{5.15}$$

5.9 Human Descriptor

In the previous sections, the extraction of *percepts* of people in the environment has been described in detail. These are the perceptual skills of a robotic system that help a robot to understand and perceive human nonverbal behaviour for the task of personality assessment. These perceptual skills are then combined together into a single descriptor P_{person} that contains all the information of interaction partner.

After a person is detected, the skeleton joint information is extracted using NiTE library (see Appendix C) and stored in the percept, $P_{skeleton}$.

$$P_{skeleton} = ((x_1, y_1, z_1), \ldots, (x_{15}, y_{15}, z_{15})) \qquad x_i, y_i, z_i \in \mathbb{R} \ ; \tag{5.16}$$

The skeleton joints are then converted into joint angles (see Section 5.3 for more details). These joint angles are stored in the following percept.

$$P_{joint\_angles} = ((x_1, y_1, z_1), \ldots, (x_{15}, y_{15}, z_{15})) \qquad x_i, y_i, z_i \in \{0 - 360°\} \ ; \tag{5.17}$$

After face detection, skeleton joint information is used to fuse the faces with the respective skeleton (see Section 5.2 for more details), and the coordinates of the face are stored in the percept, P_{face}.

$$P_{face} = (x, y, z) \qquad x, y, z \in \mathbb{R} \ ; \tag{5.18}$$

The percept, $P_{posture}$, has the information of postures, posture ID and the type of posture (see Section 5.3 for more details), as illustrated in the following equation.

$$\begin{aligned} P_{posture} = (id, p, pt) \qquad & id \in \mathbb{N} \ ; \\ p \in \{ & crossed\_arms, open\_arms, pointing, thinking\_head, relaxing, \\ & thinking\_chin, shrug, standing\_normal, aggressive\_stance\} \\ pt \in \{ & standing, sitting\} \end{aligned} \tag{5.19}$$

The percept, P_{static_hand}, has the information of hand coordinates and the gesture performed, as shown in the following equation.

$$\begin{aligned} P_{static\_hand} = (x, y, z, g) \qquad & x, y, z \in \mathbb{R} \ ; \\ & g \in \{thumb\_up, thumb\_down, okay, victory, inquire, point\_up, \\ & \quad number_1, number_2, number_3, \ldots, number_{10}\} \end{aligned} \tag{5.20}$$

Similarly, the percept, $P_{dynamic\_hand}$, has the information of dynamic hand gestures and the gestures ID, as shown in the Equation 5.21.

$$\begin{aligned} P_{dynamic\_hand} = (id, dhg) \qquad & id \in \mathbb{N} \ ; \\ & dhg \in \{no, come, crazy, blah - blah, scolding, waving, power, \\ & \quad going\_up, going\_down, clockwise\_rotation, swirling\} \end{aligned} \tag{5.21}$$

The percept, $P_{proximity}$, contains the distance information between the robot and the interaction partner, the proximity stance of the person and the change in proximity value, as given in the following equation.

$$\begin{aligned} P_{proximity} = (z, proximity\_stance, \triangle P) \qquad & \triangle P, z \in \mathbb{R} \ ; \\ proximity\_stance \in \{forward, neutral, backward\} & \end{aligned} \tag{5.22}$$

The activity information of a human is stored in the percept, $P_{body\_movements}$, which is illustrated in the Equation 5.23.

$$P_{body\_movements} = (activity) \qquad activity \in \{active, inactive\} \ ; \tag{5.23}$$

Duration of speech is also an important percept. It is stored in the percept, $P_{speech\_duration}$.

$$P_{speech\_duration} = (duration) \qquad duration \in \{long, short\} \ ; \tag{5.24}$$

Facial expression of a person is stored in the percept, $P_{expression}$. This percept also has the information of the face coordinates.

$$\begin{aligned} & P_{expression} = (x, y, z, e) \qquad x, y, z \in \mathbb{R} \ ; \\ & e \in \{neutral, angry, disgust, fear, happy, sad, surprise\} \end{aligned} \tag{5.25}$$

The percept, P_{action_units}, has the information of all the action units as shown in the Equation 5.26. Apart from some standard action units, this percept also has the information of combination of action units, such as $AU1 + AU4$.

$$\begin{aligned} P_{action\_units} = (x, y, z, au_i \forall i \in AU) \qquad & x, y, z \in \mathbb{R} \ ; \\ AU = \{ & 1, 2, 4, 5, 6, 7, 9, 10, 11, 12, 15, 16, 17, 20, 22, 23, 24, 25, 26, \\ & 27, 1 + 4, 1 + 2 + 4, 4 + 5\} \end{aligned} \tag{5.26}$$

The head pose of a person is stored in the percept, $P_{headpose}$.

$$P_{headpose} = (x, y, z, \phi, \theta, \psi) \qquad x, y, z \in \mathbb{R} \ ; \ \phi, \theta, \psi \in [-90°, +90°] \tag{5.27}$$

The head gestures are important percept. They are stored in the percept, $P_{headgesture}$.

$$\begin{aligned} &P_{headgesture} = (id, hg) \qquad id \in \mathbb{N} \ ; \\ &hg \in \{nodding, shaking, tilting, looking\} \end{aligned} \tag{5.28}$$

The *percept* P_{person} is combination of all the above perceptual abilities. These skills are combined together to construct a single descriptor, as shown in the following equation.

$$\begin{aligned} P_{person} = (&P_{skeleton}, P_{joint\_angles}, P_{face}, P_{posture}, P_{static\_hand}, P_{dynamic\_hand}, \\ &P_{proximity}, P_{body\_movements}, P_{speech\_duration}, P_{expression}, P_{action\_units}, \\ &P_{headgesture}, P_{headpose}) \end{aligned} \tag{5.29}$$

5.10 Discussion

This chapter presents the development of different perceptual skills of the robot. The chapter shed lights on the significance as well as the implementation of these skills. These skills in the context of HRI include understanding, recognition and detection of the human face, body postures, static and dynamic hand gestures, human proxemics, body movements, facial expressions and head gestures. These skills are highly important for further analysis of human behaviour to assess human personality traits.

The first step in the assessment of human personality traits is the detection of human. A person is detected either by face detection, human skeleton detection or both. To limit the computations in order to recognize all the perceptual abilities in real-time, person standing near to the robot is termed as an interaction partner and other humans present in the environment are excluded from processing. After detection and localization of human face and body, the visual perceptual system starts recognizing many nonverbal cues. Behavioural cues depict the personality traits of a person, e.g., extroverts generally have open body postures, and they are quite animated during interactions.

To detect body postures, human joints positions are converted into meaningful joint angles. These angles are, then, preprocessed and used to generate a feature vector for SVM classification. This posture module ensures that different postures that represent certain personality traits, such as dejected and self-touching postures represent neuroticism, are recognized for the latter task of personality trait assessment. Similarly, hand gestures also represent the inner emotional state of a person, e.g., thumbs up gesture shows a positive attitude, which generally extroverts also have. To detect hand gestures, SIFT features are extracted and are represented using the bag-of-feature approach. These features are also classified using SVMs.

Furthermore, human proximity with respect to the robot, body movements, facial expressions and head gestures all play an important role in the assessment of human personality.

Positive facial expressions such as happy and surprise are more often associated with extroverts while sad/fear expressions with self-touching postures and looking down head gestures are associated with introversion personality traits. Extroversion personality dimension is also correlated with closer proximity of a person and introverts tend to keep a distance from their interaction partner. Hence, these perceptual abilities are highly significant for the assessment of personality traits.

Next chapter discusses the statistical significance of these perceptual skills towards big five personality traits and also elaborate on the assessment methodology implemented in this thesis for personality traits.

6. Human Personality Traits Assessment

Knowledge of human personality is highly relevant as far as natural and efficient HRI is concerned. The idea is taken from human behaviourism, with humans behaving differently based on the personality trait of the communicating partners. However, culture, social status etc. are also contributing factors in this regard. Robots can also behave appropriately according to the human personality type. However, recognizing human personality is an extremely challenging task. Even for humans, sometimes it is difficult to know the personality of our counterparts. Numerous psychologists have presented their theories which report some features that can represent human personality type. Most of these features are not the defining features and, hence, do not point directly towards a personality type. However, some studies map different nonverbal cues to different personality types.

Although recognizing traits from nonverbal cues may not be precise, combining it by asking questions regarding different situations and analysing how the person responds, can provide us with a credible outcome of personality trait recognition. With the advent of social robots, research communities need to deal with the personality aspect in which a robot interacts with a human. Surprisingly, the question of how the robot should behave with its communicating partner still lies in the research domain. Humans get a minimum set of cues from the environment to recognize personality and act rationally accordingly. However, the inability of the robotic systems to assess personality traits leads to the irrational behaviour of robots.

Human personality traits assessment by a robot requires many perceptual abilities. These perceptual skills include the detection of human articulators, such as the face, hands, body, head, and so on, and recognition of subsequent percepts, such as postures, gestures, expressions and subtle nonverbal cues, such as proximity, speech duration and body movement of a person. According to [Jensen 16], personality traits can strongly be correlated with nonverbal cues. Although the perceptual skills, discussed in the previous chapter, are implemented based on the psychology and cognitive science studies, the significance and the role of each perceptual ability for personality trait assessment has not been extensively studied in the field of human-robot interaction.

This chapter discusses the significance of these acquired skills with respect to personality traits in order to validate the psychology claims and their theories. In addition, this chapter explains the previous technical works and systems that have been presented in the literature to analyse human personality traits. We also discuss the problems and challenges associated with these systems. Moreover, this chapter also elaborates the implementation of big five personality trait theory and the temperament framework for personality traits assessment, which is the primary interest of this thesis.

We introduce personality trait theories from psychology in section 6.1. Section 6.2 presents a literature review of technical systems for personality traits assessment. The inherent problems in these works are discussed in section 6.2. In the following next sections, we describe the implementation of personality traits assessment system and temperament framework.

6.1 Personality Trait Theories

Judging or recognizing personality trait is undoubtedly a cognitive aspect which requires intelligence. This judgmental process is extremely fuzzy as there are so many facets ingrained in every human. Interestingly, there is no quantitative standard to judge the severity of each facet. Intuitive and perceptive skills do the trick for the person judging the personality of another person. The research on personality has a strong background in psychology and communication studies. A proper understanding of the technical implementation of personality traits demands decent comprehension of personality trait theories in the first place.

There are many complex models available in psychology to recognize the personality traits of humans, as discussed in Chapter 3. Myers-Briggs Theory suggests that personality types can be recognized based on our preference to particular objects, ideas, facts etc. There are four pairs in this model, namely extroversion and introversion, sensing and intuition, thinking and feeling, judgment and perception [Myers 80]. Another theory deals with the temperament, which proposes four basic personality types, namely sanguine, choleric, melancholic and phlegmatic. These categories are named after the bodily humours, which have a relation to our bodily fluids [Eysenck 85].

However, the most dominant theory in the research of personality traits has been presented by McCrae and Costa [McCrae 99], known as the *Big Five* (BF) model. The model consists of five big dimensions, namely Extroversion, Open to new experience, Conscientiousness, Agreeableness and Neuroticism. Extroversion is more associated with activities and expressiveness. Agreeableness is more likely to have an association with being lenient, trusting, soft-hearted, generous etc. Furthermore, neuroticism has a close co-relation with temperament, self-pity, emotion, vulnerability etc. Persons with vivid imagination, creativity, curiosity, liberalism etc. are often deemed to be open to new experiences. Finally, conscientiousness is associated with punctuality, ambition, hard-working mentality etc. As far as this model is concerned, each dimension is considered to be a continuum or spectrum, in which the extremes are quite distinct. In other words, a person is placed somewhere on the continuum of each dimension based on the individual scores. Verbal and nonverbal facets are taken into account to recognize possible personality trait. Interestingly enough, a person can be placed in more than one dimension, but there is a dominant personality trait ingrained in each person [Jensen 16].

6.2 Related Work of Personality Traits Assessment

Among the *Big Five* (BF) personality types, extroversion-introversion dimension has been the most studied one since its inception. Some of the features that exploit extroversion include proximity, gesticulation, facial expressions, etc. In several research works, extroversion-introversion dimension has been considered to be a single continuum (from high extrovert to high introvert): high score in one indicates having a low score in the other. For example, extroverts score high on expressiveness. Therefore, introverts have a low score on the same facet. Individuals high on extroversion prefer to stand close to a conversation partner, and they also like to sit close to co-communicator. Individuals low on extroversion, prefer to stand and sit at a distance during a conversation. Similarly, extroversion is more related to the frequent use of rapid body movements while neuroticism correlated with self-touching behaviour. Extroversion is also associated with intense facial expressions, e.g., smiling a little more during the conversation.

Approaches that deal with nonverbal communication in order to assess human personality have been discussed in many studies [Batrinca 11] [Batrinca 12]. The approach applied in [Zen 10] uses interpersonal distances and the speed at which persons walk. The most important features have been extracted from the openSMILE [Eyben 10] to assess if an individual's behavioural pattern is associated with extroversion or conscientiousness. In the case of automatic personality detection from nonverbal behavioural cues, Batrinca et al. [Batrinca 11] have applied automatic detection of the BF personality traits in scenarios of self-presentation and employment interviews. Different features, 17 visual, 3 speech time, and 9 acoustic cues have been used for classification. The visual nonverbal cues include eye-gaze, frowning emotion, hand movements, head orientation, mouth fidgeting, and posture.

In their further work, Batrinca et al. [Batrinca 12] have explored to detect the BF personality traits in the human-computer interaction (HCI) scenario using the map task [Anderson 91]. The work aims to recognize the BF personality traits in a collaborative task setting. Features used for classification of personality traits are acoustic features, e.g., duration of speech, pitch, intensity and so on; visual features, e.g., motion vector magnitude over skin computed by using discrete cosine transform (DCT), and additional features such as the number of speaking turns by the subject. Authors have used support vector machines (SVMs) for the classification task. The results report an accuracy of greater than or equal to 70 percent for emotional stability, extroversion, and conscientiousness scales, with a fair chance of masking effect provided towards trait agreeableness. Although authors have argued having detected the BF traits accurately except for open to new experiences trait, a question that remains unanswered is how the lack of features, for instance, the distance between the speakers, detection of facial expression, and many more have been compensated.

The research by Staiano et al. [Staiano 11] stipulates the method of automatic classification of personality using visual and acoustic features corresponding to the BF traits. Low-level features include acoustic features, whereas, high-level features are social features consisting of head-pose and visual-gaze. For example, trait extroversion is best recognized and measured using the mean of 'attention received', 'attention given', 'the attention received while not speaking', 'energy in a frame', 'formant frequency', and 'spectral entropy'.

Salam et al. have conducted the most pertinent personality trait research on the automatic analysis of engagement in HRI [Salam 16]. The work aims to judge the impact of personality traits of human participants on the engagement with robots. Among the three phases of analysis, the first phase consists of data collection in the HRI triadic scenario, while the second phase includes extraction of individual and interpersonal features based upon nonverbal cues from human participants. Individual features include a histogram of gradients (HOG), a histogram of optical flow (HOF), body activity, joint speed, motion features, etc. Interpersonal features include the visual focus of attention, the global quantity of movement, relative orientation, and distance between the participants, and relative orientation with respect to the robot. The final phase includes a prediction of the level of engagement in two types of engagement, namely individual and group engagement, based upon the predicted the BF personality traits and the features extracted in phase 2. The authors have concluded that the prediction of engagement using personality traits reports better results as compared to when personality trait information is not used.

Another exciting research work, [Biel 12], selects 281 video blog videos for rating personality traits. They have used crowdsourcing as a way to obtain personality impressions from ordinary people during video-watching. Nonverbal cues have been extracted automatically from audio and video to describe vloggers' behaviour. Figure 6.1 shows the overall approach of this research work. With the help of Amazon Mechanical Turk (AMT) workers, they have conducted a crowd-sourcing personality perception annotation by using a short personality questionnaire [Gosling 03].

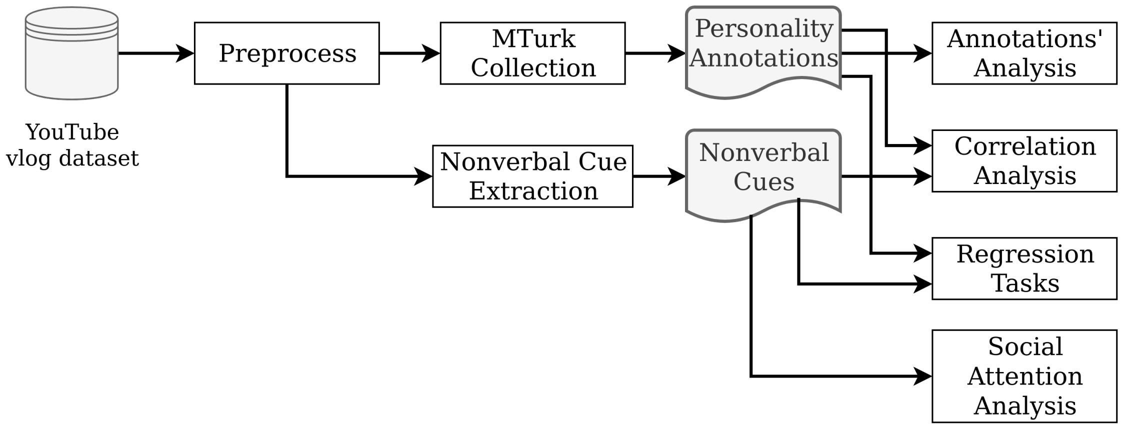


Figure 6.1: Overview of the approach for the study of personality impressions in YouTube vlogs using nonverbal cues [Biel 12].

On the contrary, Joshi et al. [Joshi 14] have used SEMAINE corpus to extract clips for personality assessment. The clips consist of 11 subjects interacting in 4 different scenarios. The work assesses BF personality traits and also assesses 4 extra traits, such as engagement, facial attractiveness, vocal attractiveness, and liability. Human experts are used to assess these personality traits and also used for the training purpose. The authors have found that controlling training data by considering experts' credibilities improves the prediction system's performance.

Another significant work conducted by Srivastava et al. [Srivastava 12] assesses big five personality traits by using visual features, i.e., facial expressions, number of faces present in the video etc., audio features, dealing with acoustic with analysis and lexical features which deals with the semantic analysis of speech. These features are combined together to construct a feature vector. They have used a regression model, Sparse and Low-rank Transformation (SLoT), and showed that the SLoT method helps on improving personality prediction accuracy. Figure 6.2 shows the algorithm of this research work. Evaluation is performed on movie characters, with answers to the questions being mapped to the personality scores.

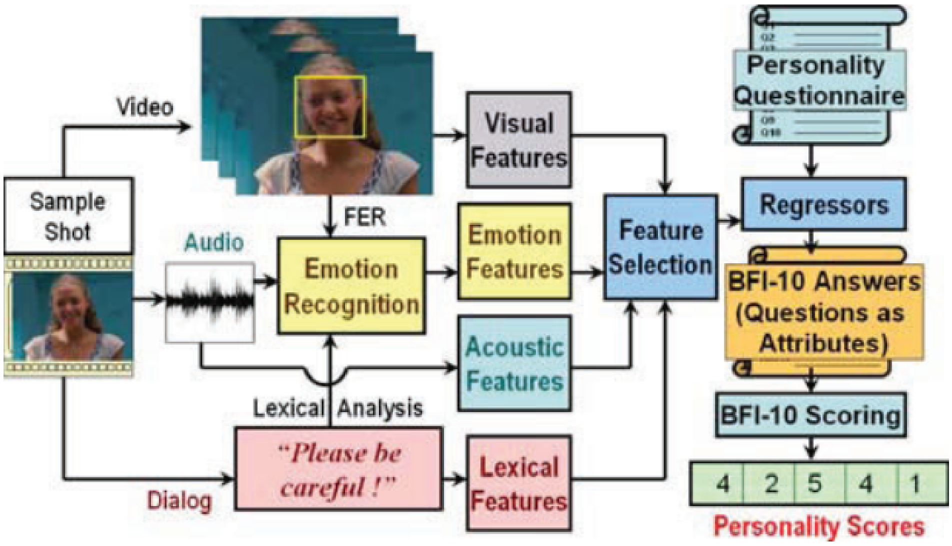


Figure 6.2: Framework proposed by [Srivastava 12] to automatically answer personality questionnaire. (Picture used from [Srivastava 12] p. 330)

Work conducted by Aly and Tapus [Aly 13] uses verbal cues and speech to extract personality traits and then using PERSONAGE natural language generator, robot adapts its speech and gestures according to the personality trait to change its overall behaviour. The authors have found that users preferred to interact more with a robot if it has the same personality type. Although the work is in the context of HRI, the approach uses naïve postures and gestures to assess behavioural traits.

Another work presented by Ge et al. [Ge 16] reports recognition of extroversion and introversion based on web browsing history and consumption records of campus cards of students. They introduce different categories, e.g., travelling, study, personal entertainment, culture and education, etc. Based on web history, features are populated in each relevant category. In addition, using the campus card of a student, they collect information about the places visited by a student, i.e., market, canteen, etc. Using Support Vector Machines (SVM) classifier, they classify extroversion and introversion with an accuracy of 72%.

An interesting approach proposed by Ferwerda et al. [Ferwerda 16] use Instagram photos to predict human personality trait. Authors extract hue, saturation and value related features and use Pleasure-Arousal-Dominance (PAD) model to get values for pleasure, arousal and dominance. The PAD emotional model describes and measures emotional states. PAD uses three numerical dimensions, namely 'pleasure', 'arousal' and 'dominance'

Table 6.1: Interpretation and summary of the correlations found between personality traits and picture properties [Ferwerda 16].

Personality	**Picture properties of the Instagram users**
Openness to experience	More green tones, lower in brightness, higher in saturation, more cold colors, fewer faces and people
Conscientiousness	Mix of saturated and unsaturated colors
Extroversion	More green and blue tones, lower in brightness, mix of saturated and unsaturated colors
Agreeableness	Fewer dark and bright areas
Neuroticism	Higher in brightness

to represent all emotions. The core idea being that physical environments influence people through their emotional impact.

However, the authors also use content-based features which estimate the number of humans present in a photo. They found Instagram picture features to be correlated with personality. A summary and interpretation of the picture features can be found in the Table 6.1. The most correlations appear in the openness to experience personality trait. However, the use of Support Vector Machine(SVM) classifier with Radial Basis Function (RBF) network has contributed to gain a recognition accuracy of 96% for an extroversion-introversion personality trait.

Work in the context of the application of personality traits has been conducted by Degroot and Gooty [DeGroot 09] to assess human personality attributions using human nonverbal cues in employment interviews. Different nonverbal cues have been considered, e.g., visual information, audio information, range of pitch, speech rate, voice breaks and fluency. Authors explore the correlation between the performance of a candidate with his/her nonverbal cues. They have found out that nonverbal cues do not necessarily lead to error in interview judgments. However, visual cues are interpreted by an expert throughout the interview.

Ventura et al. in their research work [Ventura 17] recently report an approach that uses facial information, more importantly, facial action units to train convolutional neural networks using an open-source dataset for personality traits assessment. They use Descriptor Aggregation Networks (DAN+) architecture to train the model. They explore the relationship of different facial action units with personality traits and get the overall mean accuracy of 0.912. Figure 6.3 depicts the class activation map of this approach. On the same dataset, Gorbova et al. [Gorbova 17] propose an approach in the context of *screening for job candidate*, which uses speech paralinguistic features and speech features along with facial features. A multilayer perceptron neural network has been employed for regression analysis and the fusion is performed based on the weighted sum of each feature. Their system reports 89% accuracy when tested on the dataset. However, this and similar works cover a limited range of the personality trait spectrum. In addition, the gestures and postures also express a lot about human personality which these works do not take into account.

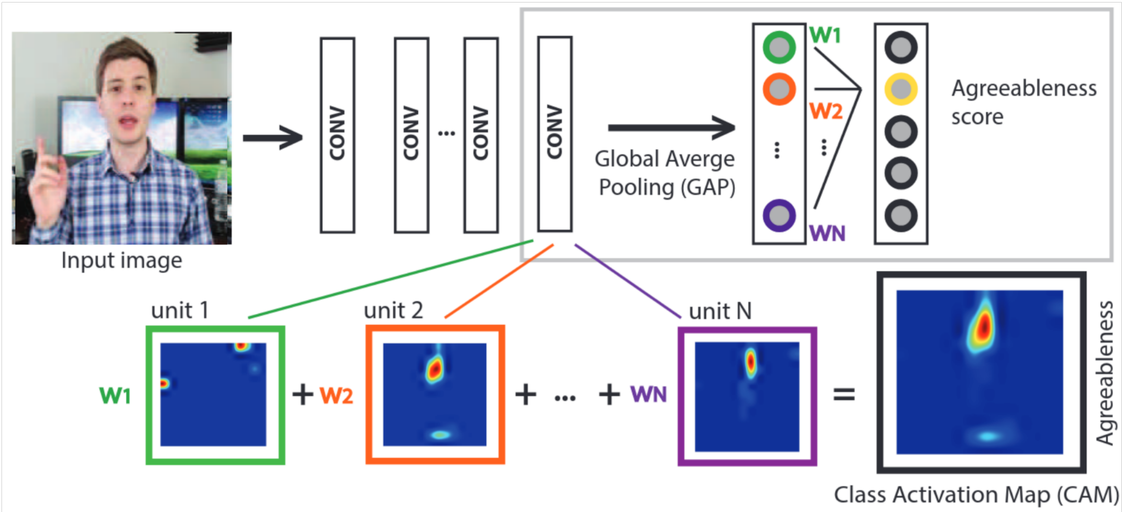


Figure 6.3: Class Activation Map [Ventura 17] scheme

The works mentioned above can assess human personality in different scenarios. However, the major shortcomings in these approaches are:

- the lack of *real-time* recognition of human personality trait based on human appearance in daily life interactive scenarios,
- the lack of availability of all the perceptual skills, such as posture analysis, proximity analysis, activity detection and gestures recognition,
- the lack of temporal analysis of human behaviour,
- the infeasibility of these systems in complex and unstructured environments,
- the lack of experimentation in the real-world environment on real-life interactive scenarios,
- the lack of assessment of subtle personality traits, such as anxiety, aggression and so on, and
- the lack of research conducted in the context of human-robot interaction to realize human personality traits assessment.

In order to address these shortcomings, the approach presented in this thesis focuses on real-time recognition of human personality traits in the context of HRI.

6.3 Big Five Personality Traits Assessment

Generally, there are two established ways in psychology, namely self-reporting and personality impressions that serve the task of personality assessment [Biel 12]. Self-reporting assessment demands that an individual judge himself based on different aspects of life. In contrast, personality impressions assessment requires a person to observe other person's personality traits. Based on the observational data provided by the observer, the personality trait of a particular person is assessed. Interestingly, studies have shown that the outcomes of both measures are quite similar [Biel 12]. However, these methods require human(s) to assess personality information. For automatic assessment of personality traits in HRI, a robot needs to have perceptual skills to recognize, detect and analyse human nonverbal cues. Nonverbal cues can be useful to bridge the gap between the technical and non-technical evaluation of personality traits. It is potentially easier to predict the behavioural pattern of a person, even in zero-acquaintance scenarios, based on the nonverbal cues.

6.3.1 Nonverbal Cues and Big Five Personality Traits

In order to understand the role and significance of nonverbal cues, a psychology survey on human behaviour analysis has been done. According to various psychologists, several human nonverbal cues indicate the emotional state as well as the personality traits of a person. Although verbal cues and paralanguage also tell a lot about the emotional state, this thesis is focused on the visual analysis of human behaviour. We describe the association of nonverbal cues with each personality trait from psychology and cognitive science studies in the following points.

- Extroverts keep their trunk wide open to people around them. It shows they are approachable to others and keeps them in a more open-minded attitude [de Vries 13].
- Individuals high on extroversion prefer to sit or stand close to the conversation partner [Knapp 13] [Hargie 16].
- Individuals low on extroversion that is introversion, prefers to stand and sit at a distance when in a conversation [Knapp 13] [Argyle 13] [Hargie 16].
- Extroversion is associated with expressiveness [Argyle 13] [de Vries 13].
- Extroversion is also associated with more facial expressions of sadness, probably as a result of their frequent social contacts [Keltner 97].
- Neuroticism & Introversion are related to low expressiveness [Argyle 13].
- Extroversion is related to more frequent and more rapid body movements [Oberzaucher 08].
- Neuroticism is involved in more self-touching behaviour [Argyle 13].
- Extroversion is associated with more frequent and more intense smiles [Ruch 94] [Argyle 13] [Oberzaucher 08].
- Neuroticism is related to facial expressions of anger, contempt and fear [Keltner 97].
- Agreeableness is positively correlated with laughter and a sympathetic facial display (eyebrows of sadness, moving the head forward and a concerned gaze) when they interact with someone in an embarrassing situation [Keltner 97].

- Extroverts typically spend more time on mutual eye contact during conversations while introverts & neurotics normally avoid mutual eye gaze [Burgoon 89] [La-France 04].
- Conscientiousness is associated with low levels of negative facial expressions and laughter. Individuals that score high on the conscientiousness scale can display a controlled smile, aversion to eye contact and touching their own face in social situations that produce some kind of distress [Keltner 97].
- People that score high on extroversion typically spend more time on mutual eye contact during conversations than other traits [Argyle 13].
- Neuroticism is negatively correlated with mutual gaze [Argyle 13] [Burgoon 89] [La-France 04] [Oberzaucher 08].
- Except for extroversion, also agreeableness and openness are associated with mutual gazing behaviour [Knapp 13].
- Extroverts talk more, they produce more words and talk longer when they have the turn [Argyle 13].
- Extroverts talk faster, louder, with shorter pauses and with a higher pitch [La-France 04] [Matsumoto 13] [Knapp 11].
- Introverts and neurotics use more and longer pauses [Argyle 13] [Hargie 16] and neurotics also produce more speech errors/hesitations [Argyle 13].
- Self-centred people convey a sense of superiority by lifting their head and tilt it backwards, and people perceive them as haughty. Similarly, they also raise their head and thrust their chin forward, and send out a ‘Don’t mess with me!’ signal [Kuhnke 12].
- Arrogant persons also come under the umbrella of self-centred personality trait. Arrogance is signalled by a slight backward tilt of the raised head and a forward thrust of the jaw [Kuhnke 12].
- Agreeable persons nod a lot during the conversation in order to encourage the interlocutor to continue [Kuhnke 12].
- Submissiveness is the sub-trait of agreeableness trait. People usually lower their heads when they are feeling submissive. Research also shows that self-touching gestures, such as holding the head at the back of the neck and placing hands on top of the head, are generally expressed in the state of submissiveness [Kuhnke 12].
- Upright posture, open arms, and a genuine smile convey ease and confidence, which is the sign of extroversion [Kuhnke 12].
- Fist slamming, sharp finger-pointing and stomping feet on the ground are all positively correlated with self-centred trait [Kuhnke 12].
- A body’s forward and erect posture with feet wide apart reveals aggression and is positively correlated with self-centred trait [Kuhnke 12].

- Open and confident posture is the norm for individuals in high-status positions such as leaders and also positively correlated with self-centred trait [Kuhnke 12].
- Introverts tend to block others by crossing their arms on their chest in a way to build a barrier. They also duck their heads and look away [Kuhnke 12].
- Similarly, neuroticism is also positively correlated with the crossing of arms on the chest. Neurotics are anxious persons with lacking in confidence and therefore, they often show self-touching postures [Kuhnke 12].

It is important to note here that humans exhibit numerous nonverbal cues during human-human interactions. The number of nonverbal cues differs significantly from one culture to the other. Past events, habitual facts and the background from which a person grows up also contribute to the type of facets he/she exhibits. The research work presented in [Kuhnke 12] provides us with an insightful explanation of how humans interpret nonverbal cues based on context or situation. Fortunately, this work also takes context and cultural information into account as far as nonverbal cues are concerned. The scope of whether a particular facet can be implemented using the existing system is also an essential factor to consider.

The capabilities of the system play a significant role to answer the question of implementability. Possible personality type is estimated based on the visual perceptual skills of the robot. Therefore, it is important to consider all those nonverbal cues which help in the accurate assessment of personality traits. For instance, the extended arm is a typical sign of openness or elation. The system has to understand this cue based on the open arms coupled with a shrug, point left or right postures. The perception of these cues makes sure that a human is showing 'extended arm'. On the contrary, backward tilt together with raised-head is a sign of arrogance, which leads to self-centred behaviour. The perception of looking up together with angry facial expressions indicates self-centred behaviour.

Sideways glance coupled with a bright smile and leaning in the forward direction, is a good indicator that a person is interested, intimate and more likely to be affectionate. Head thrust forward coupled with a red face, tight jaw, tight fist or the combination of forwarding stance, erect posture and angry face or closed palm coupled with a pointed finger are symptoms of aggression. It is a significant component of the self-centred dimension. Some signature gestures have different interpretations in different cultures. For instance, the gesture with a hand tucked into pocket or waistcoat is considered to be a sign of self-centred behaviour in some cultures. Table 6.2 shows a different type of nonverbal cues, their interpretation from psychology studies and the personality type they are associated with.

From Table 6.2, it is quite apparent that most nonverbal cues directly represent extroversion-introversion personality trait. Various nonverbal cues also represent Agreeableness-self-centred and Neuroticism-emotionally stable traits. Personality trait dimensions such as open to new experience-traditionalist and conscientiousness-careless are not directly represented by nonverbal cues. These dimensions need a lot of contextual and situational information over a long period to assess them.

Despite the reliability and diversity of the visual perceptual system that has been developed in this thesis, few nonverbal cues in Table 6.2 are not recognizable because of the technical constraints and complexity. Moreover, the system also has some technical limitations,

Table 6.2: List of nonverbal cues in relation to interpretation and personality type

Nonverbal cues	Interpretation	Personality type
Rubbing foreheads + crossed arms + holding or rubbing fingers	Self-comfort, self-touching	Neuroticism
Extended arm	Sign of openness/elation	Extroversion/ openness
Head down	Sign of being upset	Neuroticism
Upright stance + bounce in step + eyes are lively and more engaged	Sign of positivity	Extroversion/ agreeableness
Hands over the heart + tilted head + open smile	Showing appreciation	Agreeableness
Crossed legs / folded arms / a finger in front of the mouth	Sign of holding back	Introversion
Signature gesture (e.g., hand tucked into waistcoat)	Sign of pride and authority	Self-centred/ extroversion
Head hanging down + arms wrapped around body	Sign of dejection/ despair	Neuroticism
Raised arm + lightly closed fist	Sign of power	Self-centred
Backward tilt + raised head	Sign of arrogance	Self-centred
Head shaking (fast)	Sign of disagreement	Self-centred
Head nodding (slow)	Sign of agreement and encouragement	Agreeableness
Cocking one's head	Non-contact greeting	Extroversion
Self-touching + breaking eye-contact	Sign of submission	Agreeableness
Sideways glance + smile + leaning forward	Showing interest, intimacy and affection	Extroversion
Sideways glance + going away from the speaker	Lacking interest, intimacy	Introversion/ traditionalist
Tight lips	Sign of tension	Neuroticism
Blinking longer than usual	Concentration, showing interest	Extroversion
Upright posture + genuine smile + open arms	Sign of confidence	Conscientiousness
Forward facing palms + looking away	Unwillingness to get involved	Introversion
Big leaf position	Feeling of security	Emotionally stable
Palms facing downward	Dominance and control	Self-centred
Raised head / touching someone's head	Sign of authority	Self-centred
(Forward + erect posture + angry face) / (closed palm + pointed finger)	Sign of aggression	Self-centred
Forward tilting forehead + slanted gaze	Sign of disapproval	Self-centred

Table 6.3: Summary of nonverbal cues in the context of personality traits

Personality traits	**Technically Implemented Nonverbal Cues**
Extroversion	1. Open arms + shrug + point left + point right 2. Happy + proximity (near) 3. Physical activities/ movements (active) 4. Eye gaze (high) 5. Long duration of speech
Introversion	1. Sad face 2. Proximity (going far) 3. Eye gaze (very low) 4. Physically passive (less activities) 5. Block hand + not looking forward 6. Less speech duration
Neuroticism	1. Crossed arms or self-touching or think head or think chin 2. Looking down 3. Looking down + crossed arm 4. Proximity (near)
Emotionally stable	1. Looking up 2. Happy / neutral 3. Attentive (big leaf)
Agreeableness	1. Slow nodding 2. Hands on heart + head tilt (left) + happy 3. Self-touching 4. Not looking forward 5. Thinking head 6. Thinking chin 7. Crossed arms
Self-centered	1. Display of fist 2. Looking up + angry 3. Looking up 4. Standing normal + angry + point (left or right) + inquire 5. Down posture 6. Shaking head

mainly with regards to speech and verbal cues processing. The work in this thesis considers only the visual analysis of human behaviour. For the purpose of clarity, Table 6.3 shows the personality traits and the nonverbal cues that represent them.

Short speech duration coupled with low eye gaze, passivity in movement, sad facial expressions can be good indicators of introversion. On the contrary, open arms coupled with a shrug, pointing left or right, active physical movements, high rate of eye gaze and long duration of speech are viable indicators of extroversion. Agreeableness scores high on slow nodding, hands on the heart coupled with head tilting left and happy facial expressions. Thinking head or thinking chin or crossed arms are also crucial symptoms of agreeable persons. Bigleaf gesture, together with happy facial expressions is an indication of emotional stability [Kuhnke 12].

Self-centred traits have various perspectives in the way humans behave in a social setting. In some cultures, displaying fist is rude and often thought to be a sign of self-centric behaviour. This facet has a close connection with hands up. Moreover, looking up with an angry facial expression exhibited is also deemed to be a sign of arrogance, which is a dimension of self-centric behaviour. Pointing left/right or showing the inquiry sign together with angry facial expression also bears substantial evidence in favour of self-centric behaviour. To some extent, down posture using both hands indicate a sign of dominance and control. And persons with a high rate of head-shaking gestures during a typical conversation are more likely to be very self-righteous and disrespectful to another person's opinion.

In the next subsection, the methodology to assess human personality traits is discussed.

6.3.2 Methodology

A handful of technical systems have been reported in the literature for the *Big Five* personality traits assessment. Most of them either lack in considering all bodily cues or are not applicable in daily life scenarios for automatic personality analysis. As established in previous sections the significance of nonverbal cues with respect to personality, this section discusses the impact of *Big Three* personality traits on different nonverbal cues and validates the psychology claim, also discussed in [Zafar 18a]. A supervised learning strategy has been employed for this task in order to evaluate personality traits using nonverbal features. Figure 6.4 shows the working schematics of our approach.

The moment human is detected based on skeleton and face information, using the perceptual system, all the mentioned nonverbal features are extracted in real-time. Features playing no or negligible role during the traits assessment are discarded after correlation analysis. Selected features are used to train 3 different classifiers, namely extroversion-introversion, agreeableness-self-centred and neuroticism-emotionally stable in the classification stage. Recognized nonverbal cues in this thesis do not show any correlation with conscientiousness trait and openness-to-new-experiences trait. Furthermore, psychology studies have also been failed to find representative nonverbal cues for these dimensions. These dimensions require more situational and contextual cues alongside of verbal cues for their assessment. Hence, these dimensions are ignored in this thesis. SVMs are used for the classification task. A binary result from each classifier is generated, which shows whether that trait is active or not. The personality assessment methodology is discussed in the following subsections in detail.

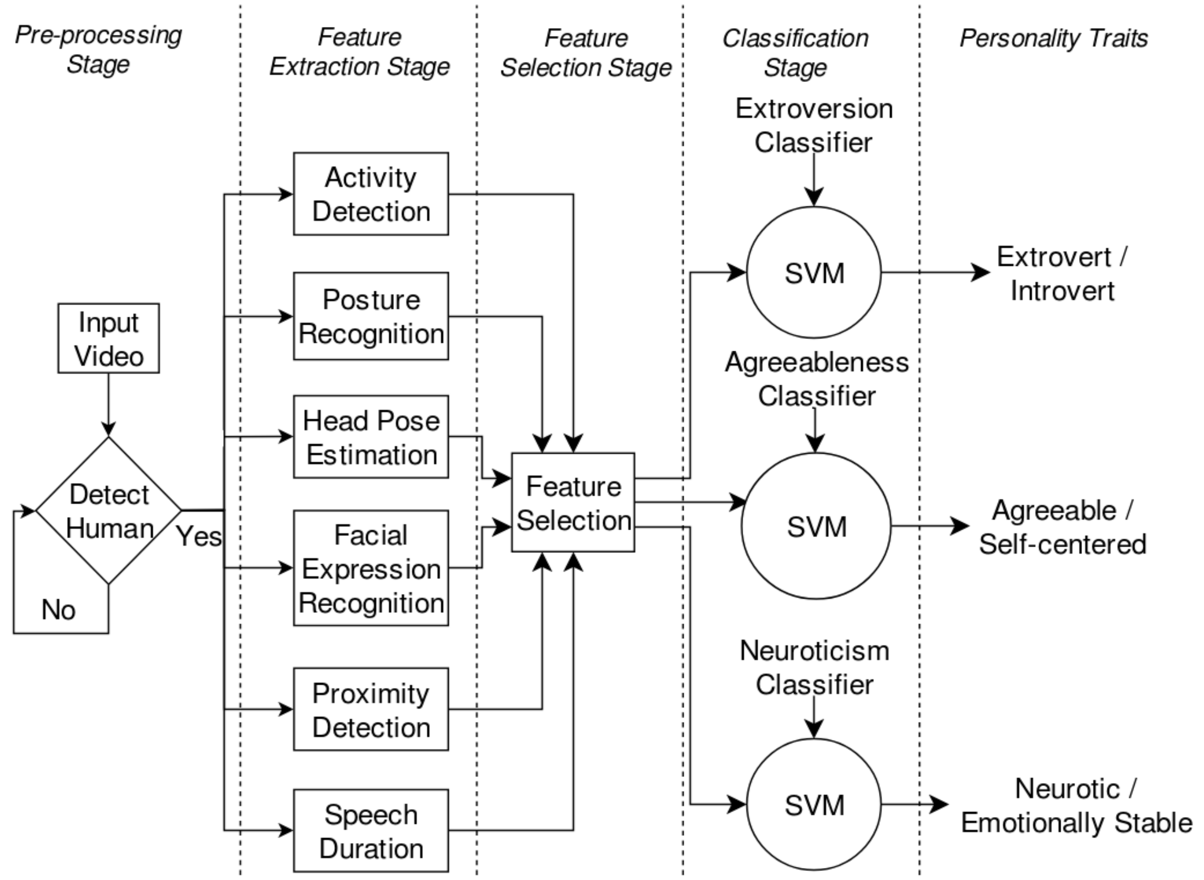


Figure 6.4: Working Schematics of BF-3 Personality Trait Assessment

Data Collection

For any classification task, having a diverse database is quite important. There are some datasets publicly available [Escalante 17] [Biel 10] [Sanchez-Cortes 12]. Most of these databases are focused on audio and video channels and don't consider depth data. Only a few uses depth information along with colour data to record human behaviour during HRI. However, these databases containing depth data are not publicly available and are not labelled according to personality traits. Due to the limitation of these existing datasets, a nonverbal feature-based database is generated.

A total of 15 participants appeared during the learning process. Each subject is asked to role-play with the robot with different personality traits. The database is focused only for 6 personality traits, namely, extroversion, introversion, agreeableness, self-centred, neuroticism and emotionally-stable personality types. The subjects have been asked to perform on different scenarios to interact spontaneously with the robot. For example, the robot, role-playing as a student, has not done the homework with the subject, role-playing as an instructor, scolding him. Each subject takes around 12 minutes to role-play for all the dimensions (half a minute session performed 4 times for each of the six mentioned personality traits).

Each session is video recorded for later analysis. All the mentioned nonverbal cues are recorded in a separate file along with the duration of speech at the end of the interaction.

These nonverbal cues, extracted every frame, are normalized and averaged over a whole sequence. The features are then used to construct a 25-dimensional feature vector. In order to label the sequences, a psychology expert is consulted. Each sequence is labelled with 3 personality traits after careful analysis of video sequences, i.e., either extrovert or introvert, agreeable or self-centred and neurotic or emotionally-stable. After preprocessing, a total of around 200 sequences are generated during this process.

Analysis of Personality Traits and Nonverbal Cues

To statistically analyse the data, correlations are used before the *learning* task. Each non-verbal cue is correlated using Pearson's correlation coefficient, as illustrate in Equation 6.1, with the traits to analyse their relationship between them and to validate the psychology claims. Negative values show that the trait is negatively correlated with nonverbal cue and vice versa. In the Equation 6.1, x and y are two variables (in our case x is a nonverbal cue and y is a personality trait), $\bar{x}$ and $\bar{y}$ are the sample mean. The coefficient value, r, can be in the range of $[-1 \text{ to } 1]$. Higher positive value of r means that the two variables are positively dependent on each other while lower negative value of r means that the two variable are inversely dependent on each other. Coefficient value close to 0 means that the variable are independent of each other.

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{6.1}$$

Table 6.4 shows correlations between different human body postures and personality traits. Table 6.5, Table 6.6 and Table 6.7 represent correlation scores for head gestures, bodily cues and proxemics, and facial expressions with 3 personality traits, respectively. We analyse the correlations trait-wise in the next sub-sections.

Table 6.4: Correlation between Personality Traits and Human Postures. (O.P. = open posture, C.P. = crossed-arm posture, C.S.= casual stance, T.P. = thinking posture, S.P. = shrug posture, P.P. = pointing posture, S.U. = stand upright posture)

Trait	O.P.	C.P.	C.S.	T.P.	S.P.	P.P.	S.U.
Extroversion	**0.80**	**-0.57**	-0.017	**-0.46**	**0.24**	0.19	0.06
Agreeableness	**-0.45**	**0.30**	0.03	**0.23**	0	**-0.33**	-0.14
Neuroticism	**-0.43**	**0.41**	-0.009	**0.41**	-0.15	-0.02	**-0.28**

Extroversion-Introversion Trait

Extroversion is associated with expressiveness. Extroverts are more likely to show physical activities during an interaction, e.g., rapid body movement, head movement, etc., which represent confidence [Jensen 16]. Correlation analysis in Table 6.6 shows that body movements play a huge role in the assessment of extroversion. They also have an open body stance which shows openness to interaction and sociability. The fact can also be seen in the interactions of celebrities in which they interact with the audience. In contrast,

Table 6.5: Correlation between Personality Traits and Head Gestures. (L.U. = looking up, L.D. = looking down, L.L.= looking left, L.R. = looking right, L.A. = looking ahead)

Trait	L.U.	L.D.	L.L.	L.R.	L.A.	Nodding	Shaking
Extroversion	0.10	**-0.38**	-0.15	0	0.4	0.16	-0.13
Agreeableness	-0.11	0.10	-0.13	0	0.1	**0.31**	**-0.27**
Neuroticism	-0.15	**0.64**	0.001	-0.04	**-0.30**	**-0.22**	0.07

Table 6.6: Correlation between Personality Traits and Bodily cues and proxemics. (C.P. = close proximity)

Trait	Body Movements	Forward Stance	Backward Stance	C.P.
Extroversion	**0.63**	0.14	**-0.28**	**0.21**
Agreeableness	**-0.52**	0.09	**-0.28**	0.15
Neuroticism	**-0.23**	**0.30**	0.10	**0.20**

introverts show more self-touching postures, e.g., crossed arms, etc. in order to block others for interaction or self-comfort themselves [Kuhnke 12] [Argyle 13]. Table 6.4 illustrates that extroversion is positively correlated with *open body* and *pointing* postures, showing the high contribution of these features towards extroversion. However, this trait is negatively correlated with *crossed-arm* and *thinking* postures, showing the high contribution of these features towards introversion.

In addition, extroverts show a varied number of facial expressions and tend to make direct eye contact with their communicating partner. On the other hand, introverts avoid mutual eye gaze and have problems in expressing emotions [Argyle 13] as can also be seen from Table 6.5. Extroverts tend to stand near the interaction partner, whereas introverts feel comfortable keeping a marginal distance [Knapp 13] [Argyle 13] [Hargie 16]. This fact can be validated from Table 6.6, which shows a backward stance is negatively correlated with this trait while forward stance and proximity close to the interaction partner has a positive correlation.

Facial expressions are equally important as any other nonverbal facet. As mentioned earlier about the role of facial expressions in the context of extroversion-introversion trait, this feature is used to analyse expressive emotions from the face. During speaking in

Table 6.7: Correlation between Personality Traits and Facial Expressions.

Trait	Happy	Sad	Surprise	Fear	Digust	Angry
Extroversion	0.16	-0.11	0.03	-0.30	0	-0.006
Agreeableness	0.12	0	0	0.17	0	**-0.20**
Neuroticism	**0.26**	0.18	0	**0.27**	-0.01	0.007

an interaction, however, this module performs poorly. The main reason is the dynamic nature of facial action units during a speech which leads to false recognition of expressions. Nonetheless, this feature has a trifling contribution towards extroversion-introversion trait. Table 6.7 shows that happiness is positively correlated with extroversion, while fear and anger are correlated with introversion.

Agreeableness - Self-centred Trait Assessment

Agreeableness trait is mainly associated with submissiveness which implies self-touching behaviour [Kuhnke 12]. Agreeable persons are more likely to be soft-hearted and generous [Keltner 97]. They tend to nod quite often during an interaction which shows their agreeable nature [Kuhnke 12] which can also be verified from Table 6.5 that head nodding is positively correlated with the agreeable trait. Moreover, they show positive and sympathetic facial expressions along with head tilt to show empathy [Keltner 97] [Kuhnke 12]. These people like to have mutual eye gaze [Knapp 13].

Self-centred persons, on the other hand, tend to look with their chin up [Kuhnke 12]. As can be seen from Table 6.5, looking up is negatively correlated with agreeableness trait. They point towards the co-speakers and express anger to show authority and shake head to show denial during communication [Kuhnke 12]. These facts can also be validated from Table 6.4 and Table 6.5 in which *pointing* posture and *shaking* head are highly and negatively correlated with agreeableness trait. Furthermore, self-centred persons also appear to stand far from the interaction partner to show superiority as compared to their counterparts [Kuhnke 12]. Table 6.6 shows positive correlation between agreeableness trait and *close proximity* which, in other words, means negative correlation between self-centred trait and *close proximity*.

Neuroticism - Emotionally Stable Trait

Neuroticism is mainly associated with depression and self-pitying. Persons having this trait are more likely to be vulnerable and emotional. They show self-touching behaviour in order to comfort themselves [Argyle 13]. This can also be observed from Table 6.4 in which neuroticism trait is highly correlated with *crossed arms* and *thinking* postures. In addition, they avoid mutual eye gaze and display dejected posture [Argyle 13] [Burgoon 89] [La-France 04] [Oberzaucher 08]. They feel comfortable looking downwards when they speak, and they mostly express anger, contempt and fear facial expressions [Keltner 97]. Table 6.5 shows high correlation between neuroticism and *looking down* gesture. In contrast, emotionally stable persons are calm and even-tempered [Jensen 16]. They mostly interact with an open upright posture. They tend to have a mutual eye gaze with the interaction partner. Moreover, they show positive facial expressions. In addition, neurotic persons tend to stand close to the interaction partner. This can be validated from the Table 6.5 that shows a negative correlation between neuroticism trait and *looking ahead* gesture.

6.3.3 Classification

Classification plays an important role in any recognition task. In order to predict personality scores for each dimension, the regression method can be used. However, there exists no dataset openly available with ground truths. In order to learn the personality traits, a database has been generated. Since personality assessment is highly a subjective process, labelling sequences with a continuous personality score accurately is highly challenging and near to impossible. Due to the lack of continuous ground truth for personality traits, only personality traits are assigned either as present or absent for the sequences and classification has been performed. As mentioned in the section 6.3.2, nonverbal cues are used to construct a feature vector. At any given time, a person possesses multiple personality traits. However, there is always a dominant trait in every human [Matthews 03].

Therefore, multiple binary classifiers are trained using SVMs for each trait. 5-fold cross-validation approach is used to optimize SVMs parameters. The Gaussian kernel has been used with 5 kernel scales. Each binary classifier uses the same 25-dimensional feature vector during training, though the classes are distinct and assigned by psychology expert according to the dimension of personality traits. A total of 205 sequences are used during the training stage.

We discuss the experiments and performance evaluation of personality trait assessment system in the next chapter.

6.4 Temperament Framework for Personality Trait Assessment

Although a limited number of technical systems have been reported in the literature, such as the one mentioned in section 6.3, for real-time personality traits assessment, these systems at best can *only* recognize the *Big Five* BF personality traits. These systems are directly based on the visual or verbal cues or both of these for accurate assessment of personality traits. They are unable to distinguish between subtle personality traits, for example, shyness and introversion or dominance and aggression. According to Watson and Clark [Watson 85], extroversion can be subdivided into the more specific facets of assertiveness, gregariousness, cheerfulness, and energy. Similarly, neuroticism can be subdivided into loneliness, anxiety, and sensitivity to rejection, while shyness is the part of introversion trait.

Before assessing these subtle personality traits, one must need to understand the concept of emotional spaces or dimensions. Like in natural sciences in which there exist dimensions such as mass, length, time, and so on, researchers also have defined few such representative dimensions in social sciences to represent human emotional states. These dimensions are commonly known as emotional space. Points in this space define individuals, segments or regions of the space define personality types, and straight lines drawn through the intersection point of three axes define various personality dimensions.

Researchers have come up with several emotional spaces. According to Wundt and Judd [Wundt 97], the three dimensions of emotions are namely, “pleasurable vs unpleasurable”, “arousing vs subduing” and “strain vs relaxation”. Many emotional spaces have been presented in psychology. Among them, the prominent ones are the circumplex model by Russell [Russell 80] and the Positive Activation-Negative Activation (PANA) model presented by Watson and Tellegen [Watson 85]. The circumplex model of affect suggests the distribution of emotion over a circular two-dimensional space, which consists of valence dimension on the x-axis and arousal on the y-axis. On the other side, the PANA model, also known as the consensual model, is known to be a 45-degree rotational version of the circumplex model. In PANA, the dimensions of arousal and valence lay at an angle of 45-degrees to the x-axis and y-axis, represented by negative activation and positive activation, respectively.

6.4.1 Implementation of P.A.D. Emotional Space

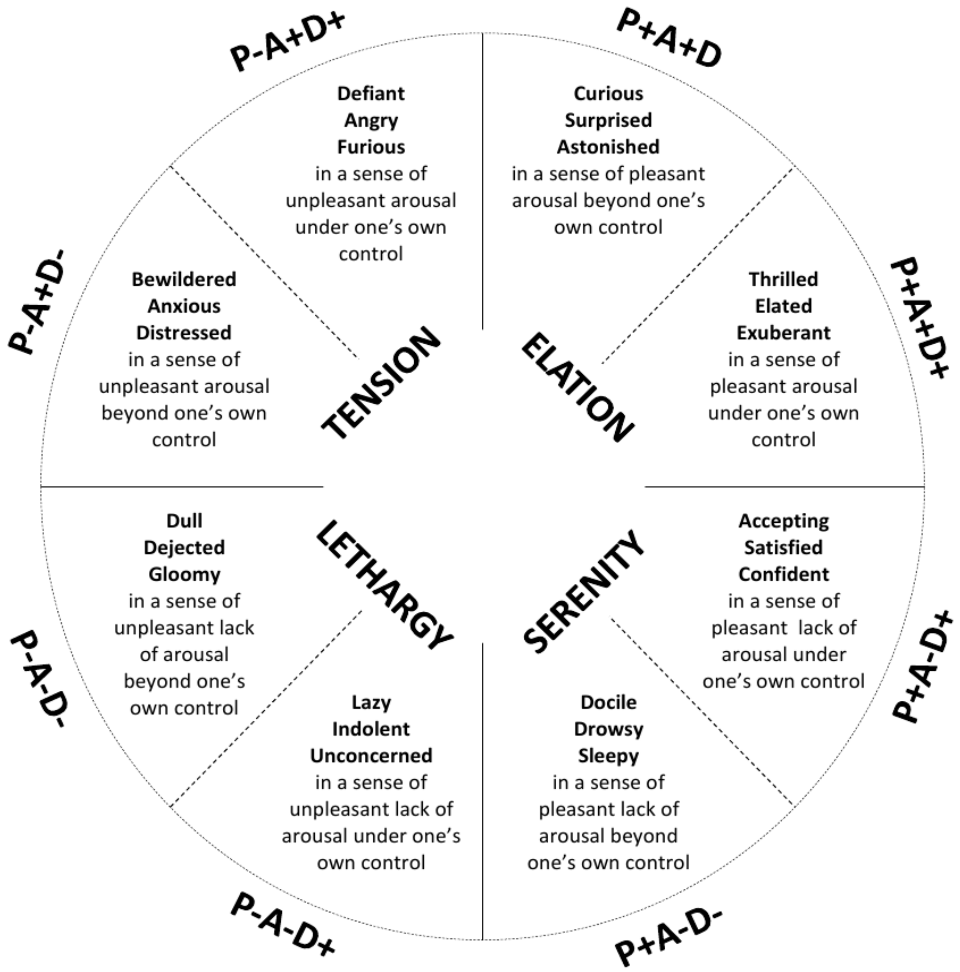


Figure 6.5: Eight affective families represented in a pie graph with 4 pleasure-arousal combinations. (Picture taken from [Boedeker 16], p. 4)

There exists another renowned three-dimensional emotional space, called Pleasure-Arousal-Dominance (P.A.D.) emotional space, presented by Russell and Mehrabian [Russell 77]. The P.A.D. model aims to describe and measure emotional traits that correspond to human personality traits. The three dimensions are defined to be bipolar such that pleasure is described as a continuum that ranges from intense pain or unhappiness on one end to intense happiness or ecstasy on the other. Arousal has been reported to range from sleepiness and drowsiness to a high level of alertness and excitement. Dominance varies from emotions of a complete absence of control or impact over events to feeling influential and in control of the situation at the opposite extreme.

With regard to human emotion, researchers have categorized emotions as either discrete and different or grouped on the basis of dimensions. In this dimensional approach, emotion serves as a point in the continuous emotional space represented by distinct dimensions that strive to realize human emotions. The model also attempts to draw a parallel with the interconnection of different emotional states based on common neural systems.

If emotions are described appropriately in terms of pleasure-displeasure, arousal-nonarousal, and dominance-submissiveness, then, identification of basic dimensions of temperament follows simply and logically [Mehrabian 78]. Temperaments can be defined as an individual's generalized emotional predisposition and be assessed in terms of characteristic patterns and/or averages of the states of pleasure, arousal, and dominance across representative life situations. This framework is based on pleasure, arousal, and dominance (P.A.D.) emotional space. After extensive research, authors have defined these three domains as follows. Pleasure can be determined using cognitive judgments of evaluation, i.e., higher evaluations of stimuli associated with greater pleasure induced by stimuli. Arousal corresponds to judgments of high-low stimulus activity using measure of stimulus "information rate". Dominance is defined as judgment of stimulus potency, with more significant the influence of stimuli corresponding to lower values of dominance.

According to the study [Mehrabian 96], P.A.D. emotional space can be divided into 8 regions based on both the extremes of each axis, denoted as P+ and P- for pleasant and unpleasant, A+ and A- for arousable and unarousable, and D+ and D- for dominant and submissive, temperament, respectively. In the Figure 6.5, the aforementioned three-dimensional P.A.D. emotional space is presented in a two-dimensional pie-type graphic, so that each of the four pleasure-arousal combinations, i.e. elation, serenity, lethargy and tension popularized by [Seo 08], are further divided according to dominance to form eight "affective" families. Each family is represented by a set of adjectives which differentiate them from other families.

The approach in this thesis is based on the work of Mehrabian [Mehrabian 96], which uses P.A.D. emotional space for personality traits assessment. In the following sections, a methodology that uses nonverbal cues for the implementation of P.A.D. emotional space is presented.

Pleasure

As previously mentioned, the value on the pleasure scale describes how much the event is enjoyable for a person. The study on facial expressions conducted by Boukricha et al. [Boukricha 09] shows that pleasure is directly associated with facial expressions. If a person is happy, the value on the pleasure scale is high. Similarly, if a person is unhappy and exhibits facial expressions such as sadness, fear, anger, or disgust, then the value on the pleasure scale is low.

To estimate the pleasure value, facial expressions percept, $P_{expression}$, which has been developed by Al-Darraji et al. [Al-Darraji 16a], as discussed in Chapter 5.6, has been used to recognize six basic facial expressions, namely happy, sad, angry, fear, disgust, and surprise in real-time. The facial expressions, extracted in every frame, are standardized and averaged over a 10 second period. Happiness and surprise expressions contribute towards the positive value of pleasure scale, while sad, fear, disgust, and anger contribute towards the negative value of pleasure scale. Algorithm 6.1 shows the estimation of pleasure value.

Algorithm 6.1: Estimation of Pleasure Value

```
pleasure = 0;
pleasant = 0;
unpleasant = 0;
average_pleasant = 0;
average_unpleasant = 0;
n ← number of frames in 10 seconds;
for i ← 0 to n do
    if Human.Face.Exist() then
        P_expression ← Human.Current.Expression;
        if P_expression = happy OR surprise then
            pleasant += 1;
        else
            unpleasant += 1;
        end
    else
        do nothing;
    end
end
average_pleasant = 1/n × pleasant;
average_unpleasant = 1/n × unpleasant;
if average_pleasant ≥ average_unpleasant then
    pleasure ← average_pleasant;
else
    pleasure ← average_unpleasant × (−1);
end
```

Equation 6.2 shows the mathematical representation of pleasure calculation. Important thing to notice is that the pleasure value is an average score over last 10 seconds.

$$Pleasure = \frac{1}{n} \max \left\{ \sum_{i=0}^{n} (H_i \vee Su_i), \sum_{i=0}^{n} (A_i \vee S_i \vee F_i \vee D_i) \right\} \tag{6.2}$$

In equation 6.2, n denotes the number of frames in 10 second duration, H = happiness, Su = surprise, A = anger, S = sad, F = fear, and D = disgust. All facial expressions are boolean and can either be active or inactive at a given time. If the second component of the equation 6.2 has a bigger value, then the pleasure score is multiplied with -1 to show the effect of negative facial expressions concerning the pleasure dimension.

Arousal

The value on the arousal scale describes how much the event is exciting and thrilling for a person. The arousal can be assessed by the combination of two nonverbal features, namely proximity and body movements. According to Nass et al. [Nass 01], people, when aroused, show frequent body movements. Similarly, Hirth et al. [Hirth 11] have established the relationship between the proximity of a person from the interlocutor and the arousal. Arousal of a person is considered high if he/she moves towards or stands close to the robot. If a person moves away or stands farther from the robot, the arousal value goes down.

Algorithm 6.2: Estimation of Arousal Value

$proximity \leftarrow [-1$ **to** $1]$ `// proximity value from percept,` $P_{proximity}$
$activity \leftarrow [-1$ **to** $1]$ `// activity value from percept,` $P_{body\_movements}$
$weight_p \leftarrow [0$ **to** $1]$ `// proximity weight`
$weight_A \leftarrow [0$ **to** $1]$ `// activity weight`
$\triangle P \leftarrow$ change in proximity `// change in proximity from percept,` $P_{proximity}$
$Arousal = 0$
$n \leftarrow$ number of frames in 10 seconds
for $i \leftarrow 0$ **to** n **do**
  if *Human.Body.Exist()* **then**
    if $\triangle P > 500mm$ **then**
      $weight_p \leftarrow 0.8$, $weight_A \leftarrow 0.2$
      $Arousal = Arousal + (weight_p \times proximity) + (weight_A \times activity)$
    else if $\triangle P > 200mm$ *AND* $\triangle P \leq 500mm$ **then**
      $weight_p = ((\triangle P - 200) * 0.00067) + 0.6$
      $weight_A = 1 - weight_p$
      $Arousal = Arousal + (weight_p \times proximity) + (weight_A \times activity)$
    else
      $weight_p \leftarrow 0.6$, $weight_A \leftarrow 0.4$
      $Arousal = Arousal + (weight_p \times proximity) + (weight_A \times activity)$
  end
end
$average\_arousal = \frac{1}{n} \times Arousal$

To calculate the arousal value, a weighted sum of proximity and body movements is used. In order to estimate the proximity value, the concept of interpersonal distances of humans during human-human interaction has been used. According to Hall [Hall 63], interpersonal distances of a person can be categorized into four zones, namely intimate space, personal space, social space, and public space. If the robot is in the public space of a person, the proximity value is negative -1. If the robot is in the intimate space of a person, the proximity value increases up to $+1$.

To detect human body movement and proximity during interaction, we use the *percepts* developed in the Chapter 5. The *percept*, $P_{body\_movements}$, contains the information about the human body movements, while the *percept*, $P_{proximity}$, contains the information of human proximity, change in proximity, $\triangle P$, and stance. Equation 6.3 shows the weighted summation of activity and proximity to estimate arousal. Algorithm 6.2 shows the estimation process of arousal while Equation 6.3 shows the calculation of arousal in a mathematical form.

$$Arousal = \frac{1}{n}\sum_{i=0}^{n}(W_P \times P + W_A \times A)$$
$$\left.\begin{matrix} W_P = 0.8 \\ W_A = 0.2 \end{matrix}\right\} \qquad \triangle P > 0.5m$$
$$\left.\begin{matrix} W_P = [0.6 - 0.8] \\ W_A = [0.4 - 0.2] \end{matrix}\right\} \qquad 0.2m < \triangle P \leq 0.5m \tag{6.3}$$

$$\left.\begin{aligned} W_P &= 0.6 \\ W_A &= 0.4 \end{aligned}\right\} \qquad\qquad \triangle\, P \leq 0.2m$$

In Equation 6.3, W_P and W_A are proximity and activity weights, respectively. The weights are dynamic and change according to the change in proximity, $\triangle P$, of a person. If a person moves more than half a meter during the interaction, the W_P gets the higher value to depict this sudden change on the arousal dimension. Moreover, the proximity value (P) is a continuous value between -1 to $+1$ and depends on how far the person is standing from the robot.

In Equation 6.3, activity, A, has a continuous value between -1 to $+1$. If a person is physically passive during the interaction, then the value is negative. The activity value becomes positive when the body movements of a person exceed the threshold value, δ. From the Algorithm 6.2 and Equation 6.3, it can be seen that W_P get more weight as compared to W_A especially in the cases where the change in proximity $\triangle P$of a person is high. The primary reason is the displacement factor. If a person A moves towards another person B and enter his/her personal or intimate space, person B would be highly aroused and quite possible act angrily. Therefore, the proximity weight, W_P, of a person has been given more weight to cover these scenarios.

Dominance

As previously stated, the value on the dominance scale represents how much the event is influential. In HRI scenario, dominance can be estimated by analysing human behaviour over time. As mentioned by de Vries et al., confident and dominant people generally have a wide-open trunk during interactions, which shows that they are approachable to others and keeps them in a more open-minded attitude [de Vries 13]. Similarly, threatening postures such as feet spread apart with hands-on-hips posture and pointing postures are correlated with aggression and dominance [Kuhnke 12]. Dominant people are also physically active during interactions [Kuhnke 12].

In order to estimate dominance, *percept* $P_{posture}$ has been used to extract different human postures, e.g., pointing posture, thinking posture, crossed arms posture, open arms posture, aggressive posture and so on. In addition, *percept* $P_{head\_gesture}$ has been used for the estimation of dominance. Head gestures such as head nodding, head shaking, look left, look down, look up, look right, look ahead, etc. are considered in this study. Equation 6.4 shows the calculation of the dominance scale. As it can be seen from the equation, open postures, pointing postures and aggressive stance depict dominant people. Similarly, mutual eye gaze with head slight upward along with active body are the other cues that show dominant people.

$$D_P = \frac{1}{n}\sum_{i=0}^{n}\{(O.P_i \vee P.P_i \vee A.S_i) \wedge (L.A_i \vee L.U_i) \wedge (B.M_i)\} \tag{6.4}$$

In Equation 6.4 and Algorithm 6.3, n is the number of frames in 10 seconds, O.P = open posture, P.P = pointing posture, A.S = Aggressive stance posture, L.A = looking ahead gesture, L.U = looking up gesture, and B.M = body movements. All the variables in Equation 6.4 are of the bool type.

In contrast, submissive people tend to look down with slumped body postures. Submissiveness is correlated with self-touching postures, such as cross arms posture or thinking

Algorithm 6.3: Estimation of Dominance Value

```
dominance = 0;
dominant = 0;
submissive = 0;
average_dominant = 0;
average_submissive = 0;
n ← number of frames in 10 seconds;
for i ← 0 to n do
    if Human.Body.Exist() then
        P_posture ← Human.Current.Posture;
        P_head_gesture ← Human.Current.Head_Gesture;
        P_body_movements ← Human.Current.Body_Movement;
        if (P_posture = O.P OR P.P OR A.S) AND (P_head_gesture = L.A OR L.U) AND
         (P_body_movements = true) then
            dominant += 1;
        else if (P_posture = C.P OR T.P) AND (P_head_gesture = L.D OR L.A OR L.L
         OR L.R) AND (P_body_movements = false) then
            submissive += 1;
        else
            do nothing;
    end
end
average_dominant = 1/n × dominant;
average_submissive = 1/n × submissive;
if average_dominant ≥ average_submissive then
    dominance ← average_dominant;
else
    dominance ← average_submissive;
end
```

postures [Argyle 13]. Submissive people also avoid mutual eye gaze and look left or right [Argyle 13] [Kuhnke 12]. They generally are passive during interactions [Kuhnke 12]. Equation 6.5 shows the calculation of the submissiveness of a person.

$$D_N = \frac{1}{n}\sum_{i=0}^{n}\{(C.P_i \vee T.P_i) \wedge (L.D_i \vee L.A_i \vee L.L_i \vee L.R_i) \wedge (\sim B.M_i)\} \tag{6.5}$$

In Equation 6.5 and Algorithm 6.3, n is the number of frames in 10 seconds, C.P = crossed arms posture, T.P = thinking posture, L.D = look down gesture, L.A = looking ahead gesture, L.L = look left gesture, L.R = look right gesture, and ∼B.M = no body movements. All the variables in Equation 6.5 are of bool type. In order to estimate the dominance value of a person during an interaction, the maximum value of Equation 6.4 and Equation 6.5 is taken as shown in Equation 6.6.

$$Dominance = \max\{D_P, D_N\} \tag{6.6}$$

If the second component of the Equation 6.6, D_N, has a bigger value then dominance score is multiplied with -1 to show the effect of submissiveness on the dominance dimension. Algorithm 6.3 shows the calculation process of dominance dimension.

6.4.2 Subtle Personality Trait Assessment Using P.A.D. Space

Visual perception of subtle personality traits in complex scenarios is a critical task for natural interaction, which even humans sometimes find it challenging to do. For a robot with a limited perception system, assessment of subtle human personality traits in real-time is an extremely challenging task. One of the reasons for the significant difference in performance between humans and robots is the use of situational and contextual cues by humans. Moreover, the robot has to compute everything in real-time, which also hampers the overall performance. Therefore, there is a need to develop a system that can assess human personality traits accurately and in real-time.

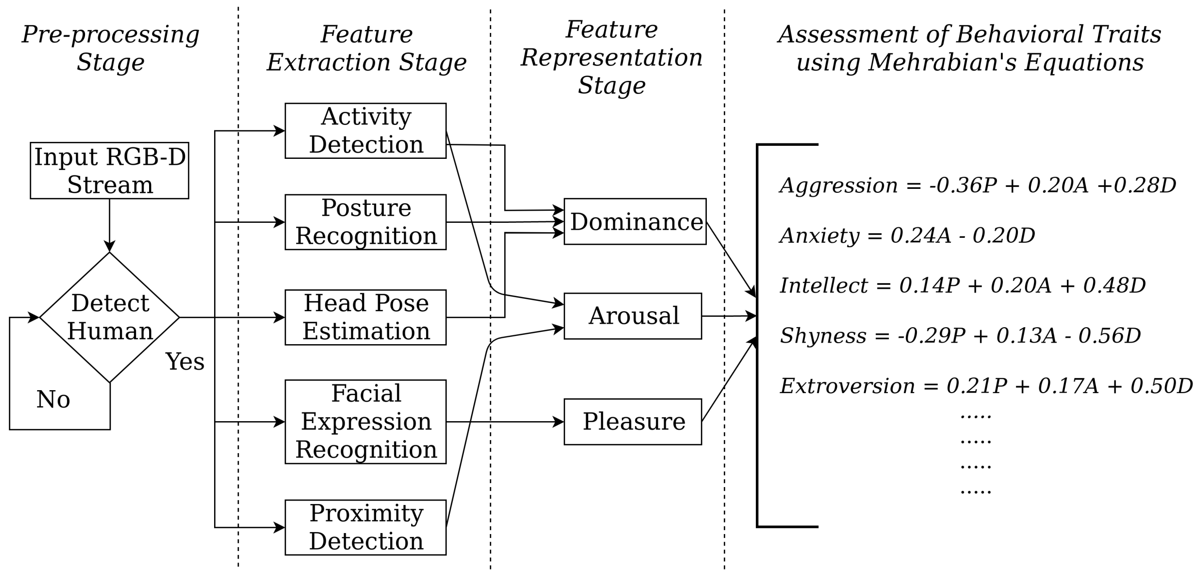


Figure 6.6: Schematic flow of subtle personality traits assessment using P.A.D. emotional space.

This thesis proposes to use the P.A.D. emotional space for the assessment of human personality traits using the Mehrabian's framework [Mehrabian 96]. Using the three dimensions, pleasure, arousal, and dominance, the author has formulated 59 individual measures that correspond to human personality traits. It has been demonstrated that traits are symmetrically related to one another based upon the P.A.D. dimensions.

Although the formulated traits are of a wide range, only 12 out of 59 traits are realized in this work. These traits are chosen according to the experimental restrictions and based on the knowledge of nonverbal cues associated with them. Personality traits such as mysticism, loneliness, and anorexic require either verbal or contextual information or both for an accurate assessment. Furthermore, even humans find it challenging to assess these traits in human-human interaction. Therefore, 12 realizable traits are considered. These traits are defined in the following, along with their equations. Figure 6.6 shows the schematic flow of the approach.

1. **Intellect**: A person who engages in critical thinking, research, and reflection about society, proposes solutions for its normative problems, and gains authority as a public figure.

$$Intellect = 0.14P + 0.20A + 0.48D \tag{6.7}$$

According to the literature [Mehrabian 96], intellect trait positively correlates with all three dimensions of P.A.D. emotional space with a stronger influence of dominance over the trait.

2. **Achievement**: Something done successfully with effort, skill, or courage.

$$Achievement = 0.13P + 0.60D \tag{6.8}$$

Achievement trait is strongly correlated with dominance and also includes secondary, but nevertheless important positive contribution from pleasure trait. Therefore, the achievement trait requires dominant characteristics that are more likely to be rewarded when accompanied by pleasantness.

3. **Extroversion**: Extroverts are behaviourally more dominant in face-to-face interactions with others.

$$Extroversion = 0.21P + 0.17A + 0.50D \tag{6.9}$$

From the equation 6.9, it can be seen that extroversion is highly correlated with dominance trait. According to [Mehrabian 71], the important behavioural factor in the study of extroversion is the postural relaxation and dominant postural behaviours. The findings have suggested that extroverts are behaviourally more dominant in face-to-face interaction with others. At the same time, they are quite expressive and social, which contributes towards the pleasure and arousal dimensions of P.A.D. emotional space.

4. **Social Desirability**: To answer questions in a manner that is viewed to be favourable by others.

$$SocialDesirability = 0.34P - 0.26A + 0.17D \tag{6.10}$$

From the above equation, it is clear that social desirability trait comes under the relaxed temperament or under serenity (see figure 6.5). Thus, the characteristic desired to be like and to make a good impression on others is associated with psychological adjustment. This is the reason why the measure of psychological maladjustment (i.e., arousal dimension) is weighted negative in the above equation.

5. **Arousal Seeking**: A person that looks for excitement, change, new environments, taking the risk, etc.

$$ArousalSeeking = 0.14P + 0.26A + 0.55D \tag{6.11}$$

This personality scale is related to exuberant temperament, but instead of characterizing specific interpersonal orientations, it tends to characterize ways in which individuals generally relate to situations. People that seek change, risk, new environments, and unusual stimuli come under this trait. This trait, therefore, is correlated strongly to dominance and arousal.

6. **Aggression**: Readiness to attack or confront someone.

$$Aggression = -0.36P + 0.20A + 0.28D \tag{6.12}$$

Emotional states such as angry, defiant, hostile, and nasty have been shown to consist of unpleasant, aroused and dominant emotional components [Mehrabian 95]. Thus, emotional traits of anger, aggression, or hostility are expected to be positively interrelated and to reflect unpleasant, aroused, and dominant P.A.D. dimensions.

7. **Dominance**: Showing power and influence over others.

$$TraitDominance = 0.72D \tag{6.13}$$

Trait dominance is correlated directly to the dominance dimension. Some researchers also claim that this trait can also have some contribution from pleasure trait. However, dominance in its purity should always represent dominance irrespective of other dimensions.

8. **Physically Active**: A person that is continuously active, working, sporting, organizing activities, etc.

$$PhysicallyActive = 0.26P + 0.40D \tag{6.14}$$

According to [Mehrabian 86], physical activity tends to be associated with more dominant and more pleasant temperament characteristics. According to the experimentation carried out in [Mehrabian 86] on individuals that participates in aerobics, running, or weight lifting, it has been found that the arousal trait has no additional impact than the general population. Rather the trait pleasure and dominance are a strong discriminator of physically active persons.

9. **Anxiety**: A feeling of worry, nervousness, or uneasiness about something with an uncertain outcome.

$$Anxiety = 0.24A - 0.20D \tag{6.15}$$

The Test Anxiety Questionnaire [Mandler 52] has been used as a means to measure Anxiety trait. The findings for the TAQ are represented in the equation mentioned above. However, this measure has been criticized by [Mehrabian 96] because of the absence of unpleasant feelings. Nonetheless, the anxiety measure employed in this work is taken from [Mandler 52].

10. **Shyness**: Nervous or timid in the company of other people.

$$Shyness = -0.29P + 0.13A - 0.56D \tag{6.16}$$

From the above equation, the shyness measure resembles anxiety trait. However, it is important to note that submissiveness is the strongest component of shyness as compared to trait anxiety. On the other hand, shyness is also an indicator of anxiety in the long run.

11. **Sensitivity to Rejection**: People who are affected easily by the negative remarks of others.

$$SensitivitytoRejection = 0.14A - 0.71D \tag{6.17}$$

According to [Mehrabian 96], sensitivity to rejection is simply a general measure of social submissiveness and understandably, a strong negative correlate of trait dominance.

12. **Nurturance**: Emotional and physical care given to someone.

$$Nurturance = 0.41P + 0.12A + 0.17D \tag{6.18}$$

Nurturance is related to giving sympathy, helping others in need, caring for children, and so on. This trait has a pattern of +P, +D, and +A coefficients decreasing in magnitudes. The expectation of pleasant and arousable characteristics for an individual with positive interpersonal are confirmed by this measure.

Using the above mentioned equations, different personality traits of a person can be estimated in the range of $[-1, +1]$ if pleasure, arousal, and dominance values are known. All these trait measures are normalized for better understanding and visualization.

6.5 Discussion

The ability of humans to adapt according to the behaviour of their interlocutor has been proven essential for an effective conversation in HHI. For social robots to interact naturally with humans, they must be well-adapted to human behaviour and personality. This chapter discusses different aspects of the personality trait assessment system. The chapter reports several technical systems presented in the literature for personality trait assessment. The challenges and the problems in these systems are highlighted. The significant shortcomings in these systems are the selection of limited nonverbal cues, lack of temporal analysis of human and lack of real-time assessment of personality traits. Moreover, these approaches are not in the context of human-robot interaction. In order to address these shortcomings, the chapter presents an approach that takes into account most of the psychology facts for personality assessment.

The chapter also highlights the importance of having perceptual abilities to recognize different nonverbal cues. The significance of these cues from a psychology perspective with respect to personality traits are discussed in this chapter. Despite the psychology claims about nonverbal cues, this thesis also validates most of those claims by correlation analysis for Big Three personality traits. An approach has been presented for the assessment of the three dimensions of *Big Five* personality traits.

Since the *Big Five* personality traits are general categories of personality types, this chapter also presents a framework to estimate *subtle* personality traits using pleasure, arousal, and dominance emotional space. The major advantage of using such a framework is that the assessment of personality traits are not directly dependent on the nonverbal cues. Instead, the assessment is done based on the pleasure, arousal and dominance values. Several ways can estimate these P.A.D. values either by verbal analysis, by nonverbal analysis or by self-report questionnaires. In this thesis, the perceptual skills of the robot are used to detect and recognize nonverbal cues in order to estimate the P.A.D. values. The framework presented by [Mehrabian 96] takes these pleasure, arousal and dominance values as input, and by using the trait equations, also presented by [Mehrabian 96], assesses the subtle personality traits. In this thesis, 12 subtle personality traits have been selected for assessment.

Next chapter discusses the experimentation and performance evaluation of personality trait assessment system.

7. Personality Trait Experiments

The implemented personality trait assessment system is organised in three separate levels, in which first level, perceptual level, is responsible for enabling the robot to perceive, recognise and understand human behaviour in the surrounding environment in order to make sense of the scenario (see Chapter 5). On the other hand, the second level, known as affective level, helps the robot to connect the knowledge acquired in the first level to make higher-order evaluations, such as assessment of human personality traits (see Chapter 6). The last level, also known as behavioural level, helps the robot in using the information from the perceptual and affective level to behave in an intelligent manner such as adapting behaviour to interlocutor mood, turn-taking, expressing empathy, and making eye contact.

The developed personality trait assessment architecture is presented in Figure 7.1. The personality system architecture has been modelled using psychology and cognitive studies. This architecture is also responsible to enable a robot to adapt its behaviour and behave in an appropriate manner.

Several individual experiments have been conducted in Chapter 5 to validate the robustness of perceptual abilities. The experiments have shown promising results for the extraction of information via nonverbal cues separately. Some part of the results have already been presented by [Zafar 16a] [Zafar 16b] [Al-Darraji 16b] [Zafar 17] and [Zafar 18c]. Since the visual perceptual system of the robot is already evaluated in Chapter 5, this chapter is focused on the interactive experiments for personality traits assessment and the performance evaluation of the proposed system.

Before the experimentation and evaluation of personality traits assessment system, we evaluate the efficiency of the implemented visual perceptual system of the robot.

7.1 Efficiency Evaluation

Detection and recognition of nonverbal cues in real-time is a critical requirement in the reliable and robust assessment of personality traits. For a social robot to show appropriate behaviour and adapt its reactions according to the personality trait of the interlocutor without any delay, the assessment of personality traits should happen in real-time. The assessment of personality traits in real-time depends upon the robust and

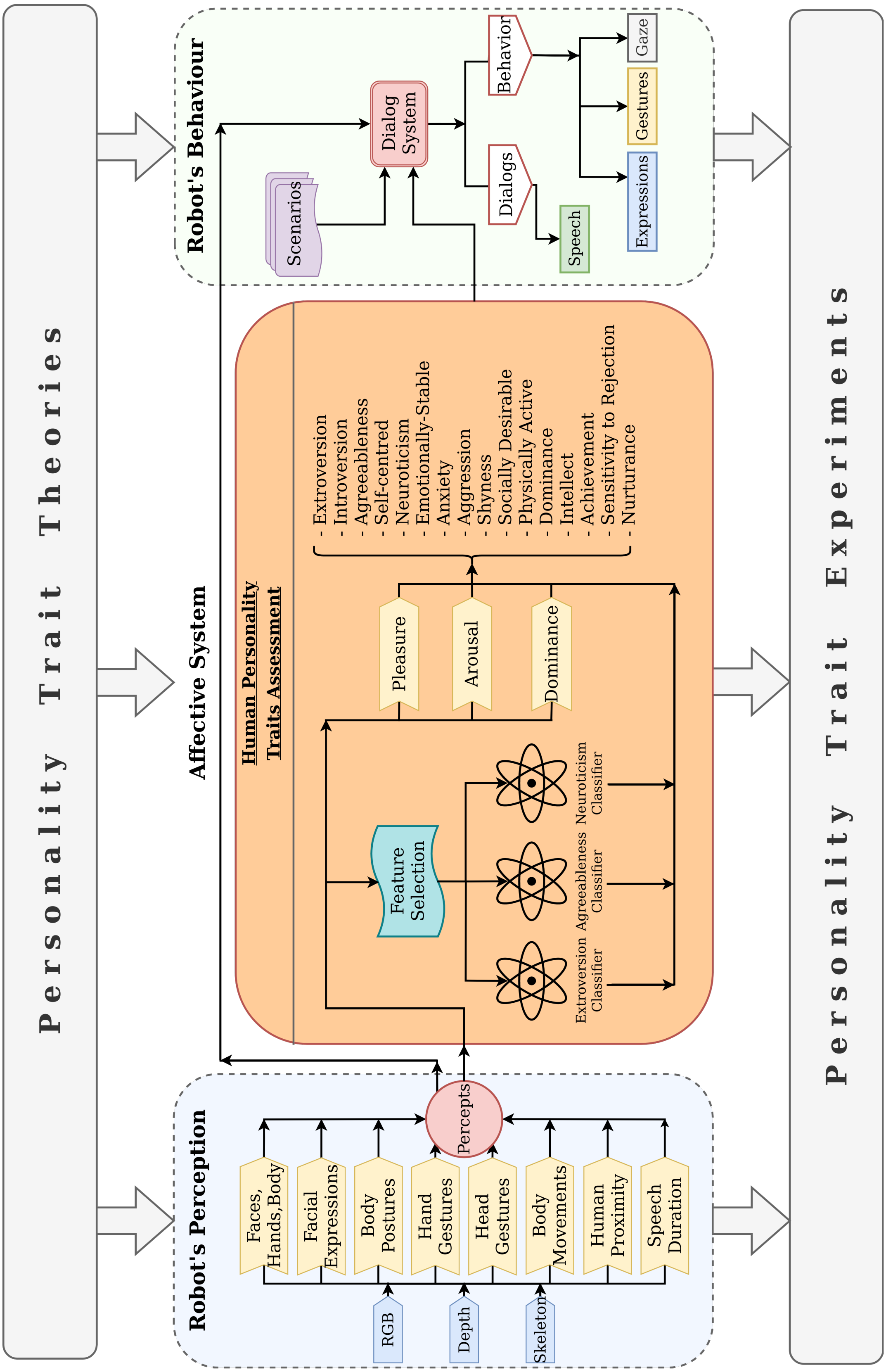


Figure 7.1: A complete overview of the developed personality trait assessment system.

efficient recognition of nonverbal features in real-time. This section analyses the processing time of all modules of the perceptual system. Table 7.1 shows the processing time of each module in milliseconds. The tests have been conducted on a system with Intel(R) Core(TM) i7-3770 CPU 3.40GHz with 16GB RAM.

Table 7.1: Efficiency analysis of perceptual system of the robot, ROBIN.

Perceptive Modules	**Average time**
Human Face Detection 2D	14.7 ms
Human Face Detection 3D	0.8 ms
Human Skeleton and Joints Information	0.003 ms
Human Joint Angles Estimation	0.009 ms
Human Posture Recognition	0.025 ms
Human Static Hand Gesture Recognition	6.1 ms
Human Dynamic Hand Gesture Recognition	56 ms
Human Proximity Analysis	0.005 ms
Human Body Movements Detection	0.012 ms
Human Speech Duration	0.006 ms
Human Facial Expression Recognition	35 ms
Human Head Gesture Recognition	2 ms
Percept Fusion	0.40 ms

From the table, it can be seen that dynamic hand gesture recognition takes the most amount of time for processing dynamic hand gestures. The reason lies in the dynamic gestures. Dynamic gestures are analysed over multiple frames to recognise them. Processing multiple frames to recognise gestures is computationally intensive. Moreover, the 2 stage classification methodology also makes the recognition process slower. However, the module still runs around 17 frames per second (FPS) and is adequate for real-time recognition. The other module that takes more time is facial expressions module. Since this module computes 23 action units values from human face, therefore, the process is computationally intensive. Nevertheless, the whole perceptual system of the robot takes around 20 frames per second to detect and recognise different nonverbal cues, which is a decent number for human-robot interaction.

7.2 Evaluation of Big Five Personality Traits

The major focus of this work is the assessment of human personality traits during human-robot interaction. Therefore, multiple scenarios have been designed in order to evaluate the performance of personality trait assessment system. The robot assesses the personality of the interlocutor and adapts its dialogues as well as its behaviour.

7.2.1 Robot Platform

To realise human-robot interaction, a humanoid robot, ROBIN, of the robotics research lab, TU Kaiserslautern has been used, as shown in Figure 7.2. ROBIN is a human-size upper body robot. It is equipped with a backlit projected face that can express more than 6 basic facial expressions. The arms make use of pneumatic muscles to show fluid motion and express human-like gestures. ROBIN also has intelligent human-like hands with 8 degrees of freedom (DoF) in fingers for flexion, the finger spread, thumb pitch, and thumb roll. There are 2 DoF in the wrist (pitch and roll), 1 DoF in elbow (pitch) and 3 DoF in the shoulder (pitch, roll, and yaw). ROBIN's compliant neck and torso both have 3 DoF. Pneumatic muscles in the arms are powered by external air supply.

For the perception task, an RGB-D sensor, Asus Xtion Pro, is installed on the chest of ROBIN. ROBIN also has an onboard pc that is responsible for control movements and expressions. A stand-alone Intel Core i7 running at 3.40GHz with 16GB RAM has been used to process the RGB-D data (see more details in Appendix A).

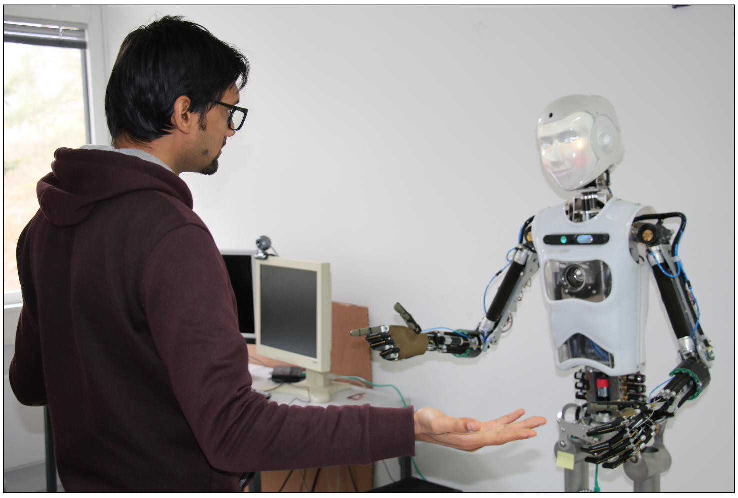


Figure 7.2: Humanoid robot ROBIN.

The robotic framework, known as *Finroc* [Reichardt 13], has been used for the development of all the perception and control modules in C++, as discussed in Appendix B. A dynamic dialogue system guides the interaction between humans and robots. This dialogue system has been developed by Koch et al. [Koch 07]. The system reads the dialogue file, which describes the flow of dialogue and builds a finite state machine (FSM). The states of the FSM represent a set of activities that the robot must do. State transitions are triggered either by a sensor input from the perception system or as a result of a callback from the application (see Appendix D for more details).

7.2.2 Experimental Setup

20 subjects have participated during the experimentation. In order to exploit the personalities of the subjects appeared in the experiments, they are provided with 3 different hypothetical scenarios to perform. First scenario concerns with two friends, one is ROBIN and other is a student, discussing the final grade of an oral exam in which the student gets a poor grade. ROBIN acts as an observer and monitors the person's behavioural traits over time. The second scenario is similar to the first one, with the student getting a good grade in an exam. The last scenario deals with an interaction between a boss and a worker in which the boss is not happy with the progress of the worker. In this setting, ROBIN

Table 7.2: Dialogues during HRI for scenario where a student gets a poor grade in oral exam

ID	Dialogues of ROBIN during Interaction
Robin	Good Morning buddy, come here. *after the subject comes closer*
Robin	How are you doing? I remember that you have an exam today? Tell me how did it go?
Subject	*Usually subject is in a bad emotional state. The subject complains about the exam and the grade. Some subjects also suggested that the marking of the exam was biased. However, some also accept the fate and motivate themselves claiming that they would work hard next time.* *Randomly any one of the next 4 dialogues can be executed given that the assessed personality of a person matches with the following condition.*
Robin-1	*If the person is found to be neurotic or introvert, robin respond in this way:* My friend you don't need to worry too much. Instead, I can help you with the preparation of the exam.
Robin-2	*If the person is found to be extrovert or emotionally stable, robin respond in this way:* At least you are still cheerful and in high spirits. Better luck next time.
Robin-3	*If the person is found to be self-centred, robin respond in this way:* Hey, you don't need to be aggressive about all this. It will be better if you put your efforts in studying hard.
Robin-4	*If the person is found to be agreeable, robin respond in this way:* This is unfortunate to hear. The good thing is that you understand where you did wrong. Put more efforts and work hard; you will get a good grade in the exam. The interaction is ended here

acts as a worker, and the person acts as a boss. A total of 37 sequences are generated during the validation stage. Each subject performs any of the two scenarios out of the scenarios mentioned. After preprocessing, three samples are removed due to failure in video recording.

7.2.3 Performance Evaluation

Table 7.3 shows confusion matrix for all the recognized personality traits. It can be observed from the table that the traits, in general, are recognised correctly. In extroversion-introversion trait, 2 sequences are wrongly predicted to be extrovert. After analysis, it has been found out that the subject in the first sequence is initially quite passive. The subject shows less body movements and is quite static which are the sign of introversion.

However, after few seconds he becomes active and dynamic. He uses his body and hands while talking. Since the personality is assessed based on the behaviour during the whole session, the personality traits recogniser assesses the person as an extrovert. An expert labels the ground truths for these sequences. Due to the subjective nature of this labelling

Table 7.3: Confusion Matrix for 3 recognized personality types (Extroversion, Agreeableness, Neuroticism).

	Predicted Extrovert	Predicted Introvert	**Recall**
Actual Extrovert	**1.00**	0.00	1.00
Actual Introvert	0.10	**0.90**	0.90
Precision	0.89	1.00	

	Predicted Agreeable	Predicted Self-centred	**Recall**
Actual Agreeable	**0.87**	0.13	0.87
Actual Self-centred	0.00	**1.00**	1.00
Precision	1.00	0.823	

	Predicted Neurotic	Predicted Emotionally Stable	**Recall**
Actual Neurotic	**0.91**	0.09	0.91
Actual Emo. Stable	0.00	**1.00**	1.00
Precision	1.00	0.96	

process, the subject could also be labelled as an extrovert. However, due to the availability of verbal and contextual cues, the human expert consulted in this work has labelled this subject as an introvert. This shows that the system can predict stereotypical personality traits accurately.

In the case of the second sequence, posture recogniser reports erroneous results due to inaccuracy of human skeleton tracker. The assessment system assumes that skeleton tracker is accurately tracking a person. However, the tracker is unable to accurately tracks the body parts of the subject. This can happen due to the following reasons.

- The person limb(s) is/are occluded.
- Person's hands are too close to the body. The tracker is unable to distinguish between the torso and the limbs, and considers them as one whole object and erroneously fits the skeleton over it.
- Full length of human is not visible. It happens when the person comes too close to the robot.

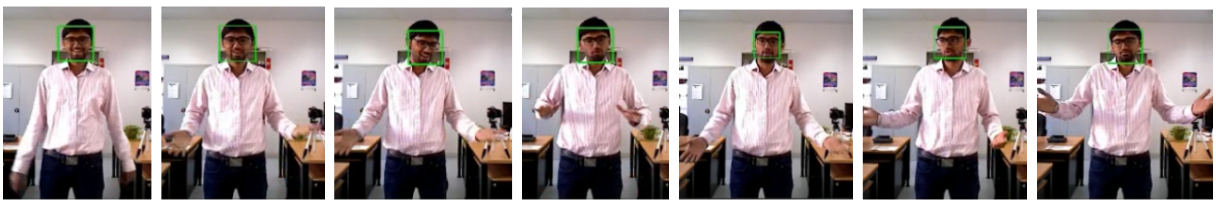


Figure 7.3: Image sequence captured during personality trait assessment with the person interacting with the robot. The person is assessed as extrovert and an emotionally stable person.

According to Table 7.3 for agreeableness personality type, 3 sequences are found out to be false positive for the self-centred trait. This trait is highly dependent on activity, postures and facial expressions. Generally, self-centred persons are quite dynamic and aggressive. In contrast, agreeable persons are soft-hearted and avoid impulsive actions/movements. Upon analysis, it has been found that the system associates activity with self-centred trait which does not hold in all situations, e.g., an agreeable person explaining a concept by stretching its arms and body.

From Table 7.3, it can be seen that neuroticism-emotionally-stable trait is recognised quite accurately. Only one instance has been predicted as neurotic wrongly. Generally, persons who are recognised as a neurotic, exhibit worried gestures and self-touching postures. They also shake their head as showing a conflict with their inner thoughts. Moreover, most of them are found to be looking down with dejected postures.

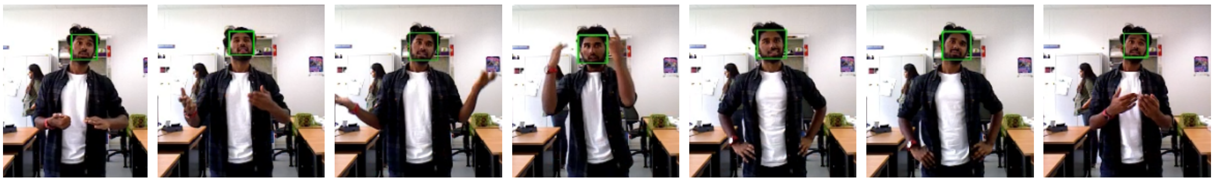


Figure 7.4: Image sequence captured during personality trait assessment with the person interacting with the robot. The person is assessed as self-centred and an extrovert person.

Figure 7.3 shows the images captured at different timestamps of a person interacting with a robot. The person is assessed as an extrovert which is also quite evident from the images. The person seems to be active with open postures and smiling most of the time. Figure 7.4 shows the images captured at different timestamps of a person in which he is assessed as a self-centred and an extrovert person. The person seems to be active with open and pointing postures. He also exhibits an aggressive stance and tilts his face upright to show dominance.

Figure 7.5 shows the images captured at different timestamps of a person in which he is assessed as a neurotic and an introvert person. The person seems to be worried with a lot of thinking and dejected postures. He also exhibits crossed arms postures in order to avoid the interaction. He also looks down most of the times, which is one of the signs of neuroticism. Overall, the system shows high accuracy for all three trait dimensions with extroversion-introversion, agreeableness-self-centred and neuroticism-emotionally-stable resulting in an accuracy of 94.6%, 91.9% and 97.3%, respectively.

Figure 7.5: Image sequence captured during personality trait assessment with the person interacting with the robot. The person is assessed as neurotic and an introvert person.

7.3 Evaluation of Subtle Personality Traits

The main hypothesis proposed is that the use of P.A.D. emotional space, computed through human nonverbal cues, can provide a successful assessment of human personality traits in the context of HRI. Due to the unavailability of ground truth, the validation of the hypothesis is a challenging task. In order to validate the authenticity of the proposed system, written feedbacks have been compiled in the form of a questionnaire from psychology students. In the following sub-sections, the experimentation and evaluation procedure is described in detail.

7.3.1 Experimental Setup

For experimentation purposes, 15 university students (12 males, 3 females; age range 22-45 years) from different ethnic backgrounds have participated. The participants have been naive about the objective and nature of the experiments, who have volunteered to participate in this study. The participants provide informed consent with regards to the guidelines of an anonymous research group.

Laboratory experiments have been conducted in a closed environment, such as an office room. ROBIN stands in front of the wall from where it can visualise the entire room, as shown in Figure 7.6. Participants have been instructed to stand in the line of sight of ROBIN. Three cameras are placed at different locations to record the interaction. ROBIN perception GUI is also recorded for the duration of the interaction. Artificial lights are used during the experiments to make it consistent for all the participants. An experimenter is also present in the room to monitor the processes systematically taking place and only intervene if the system malfunctions because of technical issues.

Figure 7.6: Top view of laboratory setup in simulation. ROBIN is standing in front of the wall.

7.3.2 Experimental Scenarios

In the direction of validating the proposed hypothesis, participants are assessed for personality traits in three different scenarios. These scenarios have been developed with regard to a student's area of exposure that consequently leads to three relatable tasks for experimentation. Each participant has been instructed to enact the following scenarios one at a time. The experimenter also introduces participants with ROBIN at the beginning of the experiments to get familiarised with ROBIN.

The first scenario involves an interaction between ROBIN, role-playing as the professor, and the participant, role-playing as a researcher. The second scenario involves an interaction between ROBIN, role-playing as a master student, and the participant, role-playing as a supervisor. The last scenario involves an interaction between ROBIN, role-playing as an interviewer, and the participant, role-playing as a candidate. For each scenario, the robot takes the lead by asking questions and responding generically. The excerpts of the scenarios are furnished below.

Scenario-1: *"You are a research assistant at a research lab. Last month you were assigned to program an important module for a project which is to be demonstrated to the group today in an hour. You have not completed the task due to several reasons. You are called into the professor's office, and you have to explain your reasons for the failure of task completion"*. Table 7.4 shows the dialogues of the first scenario during the interaction (see Appendix D, Dialogue D.1 for detailed flow of the dialogue).

Table 7.4: Dialogues during HRI for Scenario-1

ID	Dialogues of ROBIN during Interaction
Robin	Hello, come here. We need to talk. *after the subject comes closer*
Robin	I gave you a month with clear directions to finish the task. All you had to do was to spend quality time and effort to complete the task. Why is it so hard to finish on time?
Subject	*Usually subject give excuses or apologise depending on their personality.*
Robin	But this is unacceptable. If you had issues, it was your responsibility to convey them to your seniors. Are you telling me you tried your best, but still, you could not complete the task?
Subject	*Subject explain his/her reasons. Some subjects get angry, some intimidated and some don't care*
Robin	You know, you used to be one of our sharpest employees, and I was even planning on giving you a better position if you had performed well. Now all I hear is how much time you spend shopping online and chatting on Facebook. Do you have any explanation?
Subject	*Some subjects defend themselves frantically, and others accept their faults*
Robin	Okay that is enough. Go and finish the task. And let me be very clear; next time, I do not want to hear any excuses. Bye.

Scenario-2: *"You are a supervisor at a research lab in the university, and you are mentoring a master's student, ROBIN, with his thesis for the past 6 months. ROBIN, with your help, was able to obtain some good results and analysis. Due to some personal reasons, ROBIN went away for a month. Meanwhile, you published his work in a Machine Learning Conference without mentioning him. ROBIN comes back and requires an explanation".*

Scenario-3: *"You have applied for a research assistant position at a research lab. The position advertised is related to technical video editing. Today is the interview day, and you are looking forward to being at your best and get this position. ROBIN is the interviewer".*

7.3.3 Personality Traits Analysis

Figure 7.7 shows a female subject is interacting with the robot on a Professor-Researcher scenario presented in Table 7.4. During the first 10 seconds of the interaction, the subject seems quite worried. The subject has mostly negative facial expressions (i.e., sad, fear and angry) during this period which goes according to the context of the scenario.

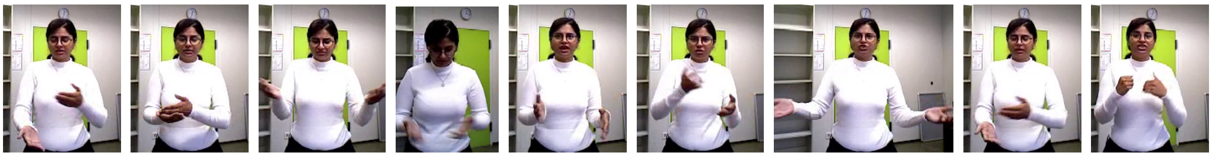


Figure 7.7: Sequence of a subject during Professor-Researcher Scenario (see Table 7.4 for detail dialogues)

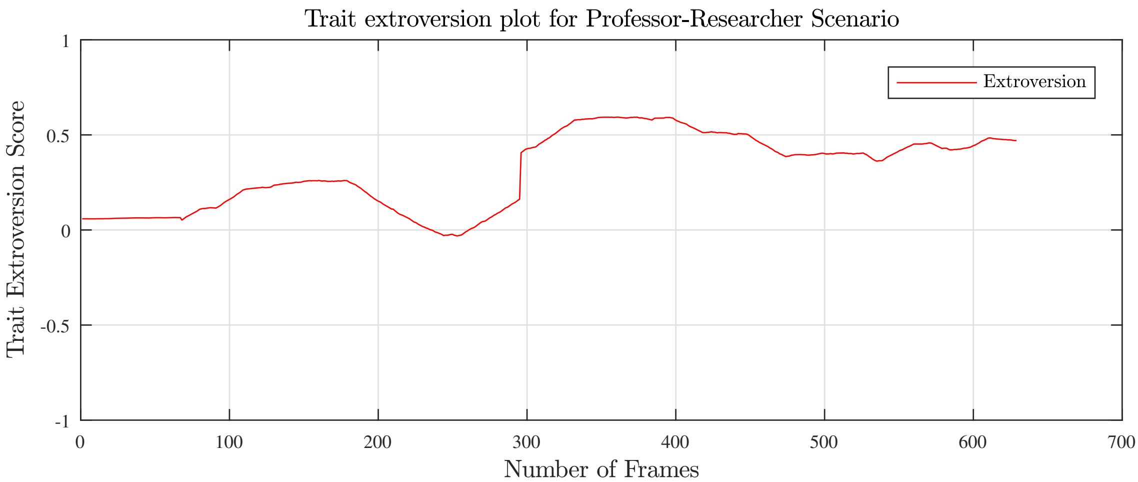


Figure 7.8: Trait extroversion plot for Professor-Researcher Scenario.

Although subject seems fairly active from the sequence, however, the subject is marginally active while thinking about the reasons why she didn't complete the assigned task. It also seems that the subject is quite intimidated by the angry robot. However, after 10 seconds the subject gets more comfortable and explains her reasons. She at times is highly active. Moreover, the pleasure value also becomes positive after 10 seconds. The subject also moves forward two times during the interaction. Subject uses her body gestures extensively to convince the robot (i.e., acting as a professor) with her genuine reasons.

Figure 7.8 shows the trait extroversion score for professor-researcher scenario, as described in Table 7.4. It can be seen from the figure, the trait score is close to zero in the start. Since the pleasure value is negative in the start, the trait extroversion score is not highly positive but nevertheless, is positive.

The reason for this positive score can be better understood from the trait extroversion equation (see Equation 6.9). Extroversion score is mostly dependent on dominance dimension of P.A.D. emotional space. Therefore, despite the negative value of pleasure, the positive value of the dominance dimension makes the extroversion score above 0. It can be seen when the subject is looking down and avoiding mutual eye gaze, the extroversion score drops at that point. Since looking down and avoiding eye gaze are the signs of submissiveness, the dominance value gets low at that particular period. After 10 second period the subject seems quite active and focused; therefore, the extroversion score is around 0.5.

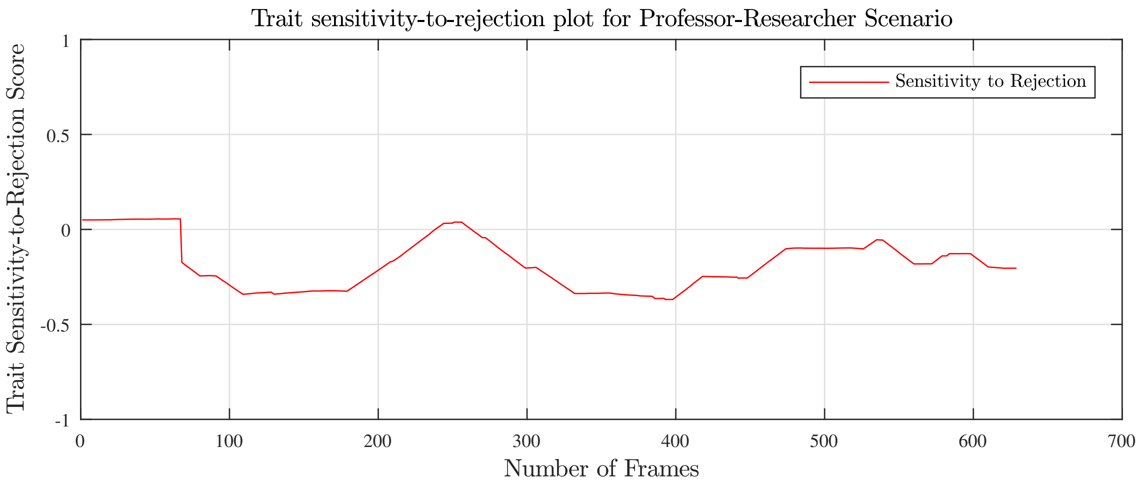


Figure 7.9: Trait sensitivity to rejection plot for Professor-Researcher Scenario.

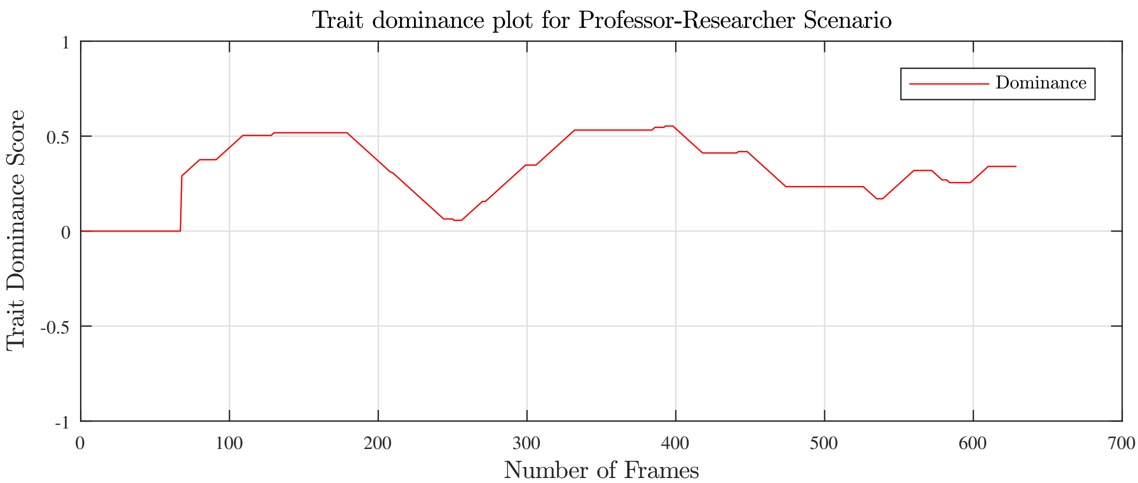


Figure 7.10: Trait dominance plot for Professor-Researcher Scenario.

Figure 7.9 shows the trait sensitivity-to-rejection score. According to [Mehrabian 96], sensitivity to rejection is simply a general measure of social submissiveness and understandably, a strong negative correlate of trait dominance. Generally, the subject during the professor-researcher scenario is fairly dominant overall. Although she seems a little disturbed in the start from the angry professor, it does not effect her attitude to defend herself. Therefore, the graphical plot of this trait is all the time negative. This fact can also be validated from Figure 7.10. As it can be seen that the subject is almost all the time dominant with just for 2-3 seconds the dominance value goes close to 0.

Figure 7.11 shows the trait shyness score. Trait shyness depends mostly on the submissiveness (opposite of dominance) trait along with unpleasant feelings. As can be seen from the Figure 7.11, the subject shyness score is close to 0 during the first 10 seconds and later becomes negative since the subject is highly dominant. Since shyness trait is also dependent on displeasure (unpleasant feelings), it can be seen that between $200-300$ frames, the shyness value becomes positive due to the negative pleasure score. However, the shyness trait reports negative score for the later half of the interaction.

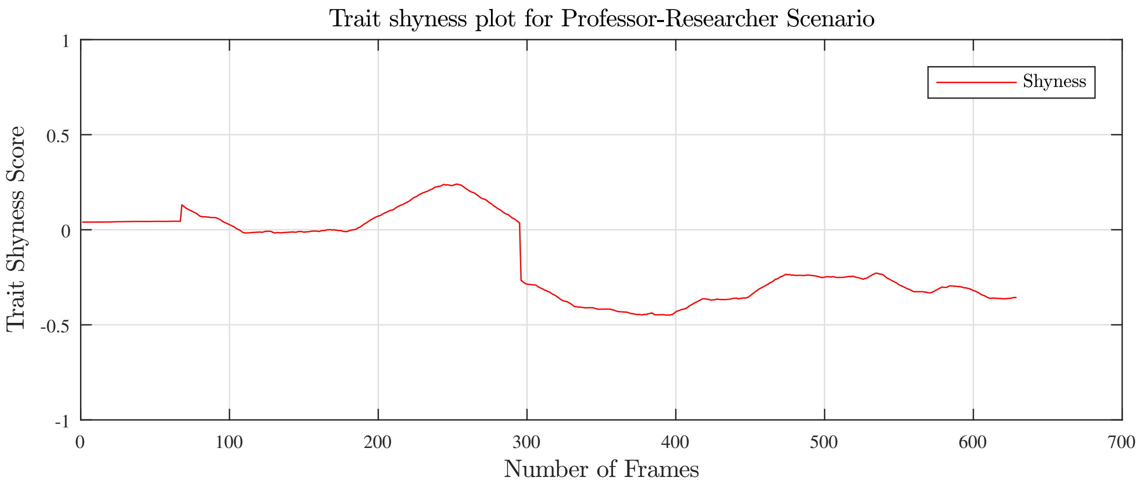


Figure 7.11: Trait shyness plot for Professor-Researcher Scenario.

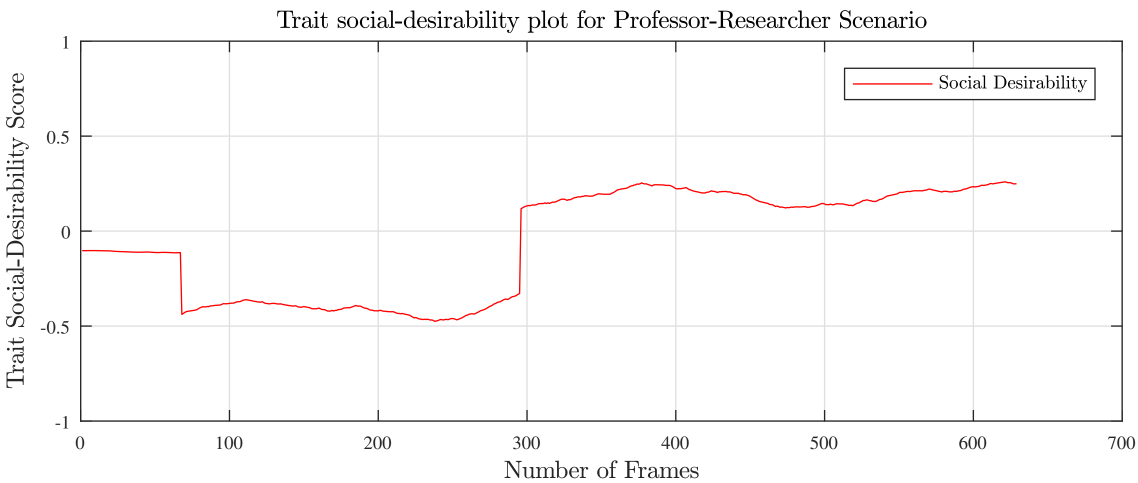


Figure 7.12: Trait social desirability plot for Professor-Researcher Scenario.

Figure 7.12 shows the trait social-desirability score. This trait is dependent on the positive value of pleasure, negative value of arousal and positive value of dominance. During the first 10*sec* of the interaction, the subject pleasure value is negative and arousal is positive. The subject is a little disturbed in the start. Therefore, the score during this period is fairly negative. However, as soon as the subject gets comfortable and the pleasure value becomes positive, the trait score also becomes positive. The important thing to notice is that the positive value is around 0.25 for this trait in the second half of the interaction. The reason for not high score is because of the fact the subject uses her body and limbs rapidly, thus, exhibiting a bit of uneasiness.

Figure 7.13 shows the trait anxiety score. At some moments during the interaction, the subject seems she has anxiety. However, overall the subject does not exhibit clearly the signs of anxiety. Since the trait is dependent on positive arousal and negative dominance, the score for this trait is positive. However, the average score of this trait during the whole interaction is ≈ 0.1, which is negligible.

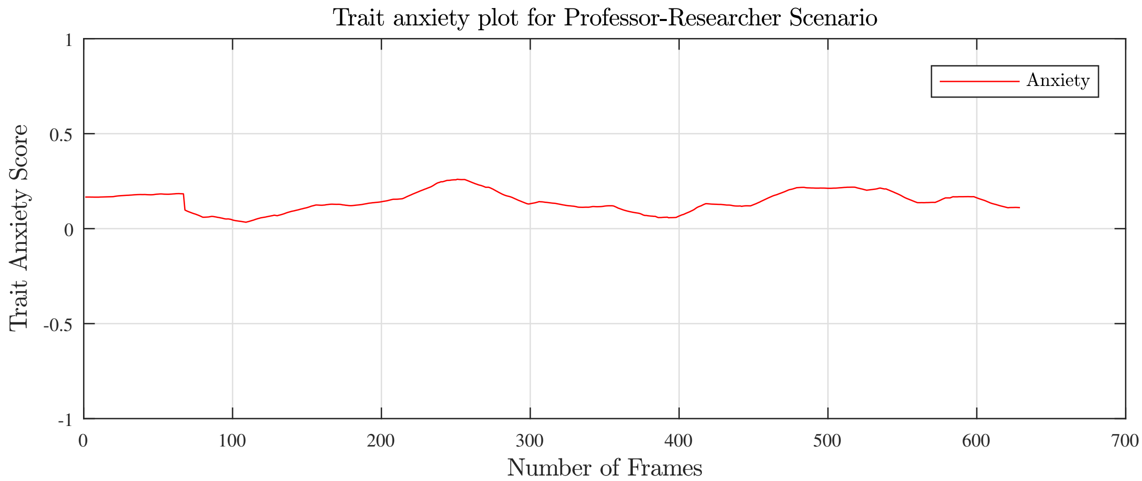


Figure 7.13: Trait anxiety plot for Professor-Researcher Scenario.

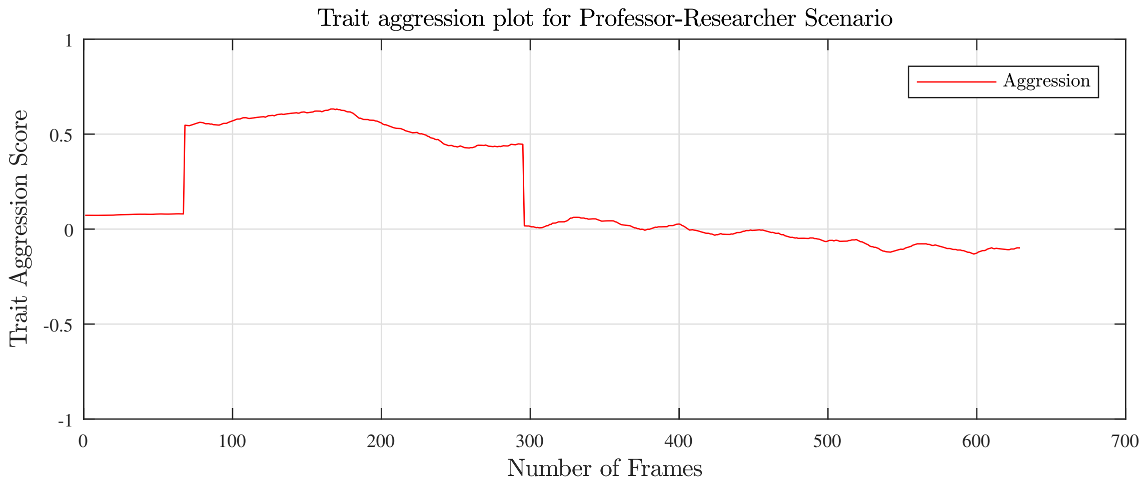


Figure 7.14: Trait aggression plot for Professor-Researcher Scenario.

Figure 7.14 shows the trait aggression plot for professor-researcher scenario. Aggression is dependent on high arousal and dominance, and negative pleasure dimensions. Since the subject has negative pleasure score during the first 10*seconds*, the aggression value is higher in that period. However, later the aggression is close to 0 value during the interaction. According to the observer analysis, the subject seems a little bit aggressive during the whole interaction. Therefore, the outcome of the system for this trait is somewhat correct during professor-researcher scenario.

Figure 7.15 shows the trait arousal-seeking plot for professor-researcher scenario. As the trait is highly dependent on the dominance value and also positively dependent on arousal and pleasure dimensions, therefore this trait is fairly positive throughout the interaction. People that seek change, risk, new environments, and unusual stimuli come under this trait. The subject during the interaction is explaining her reasons for not completing the task. The subject seems like a risk taker and is arguing strongly with the professor. Therefore, the finding of the system for this trait is fairly accurate.

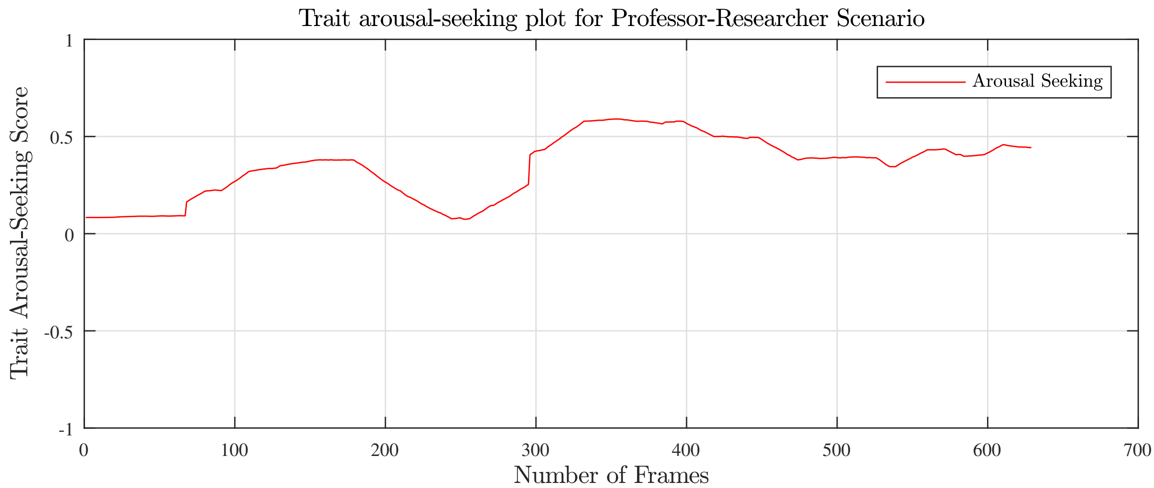


Figure 7.15: Trait arousal seeking plot for Professor-Researcher Scenario.

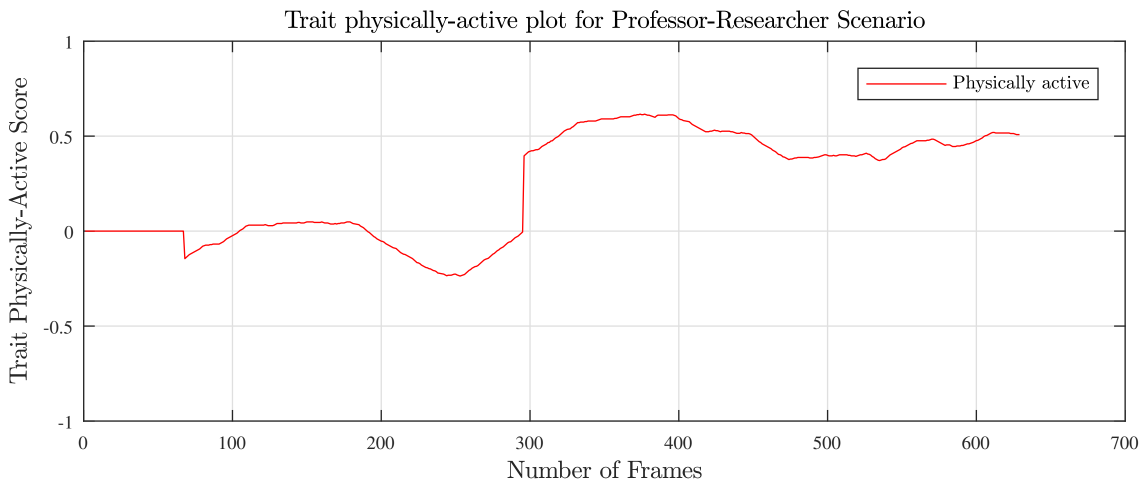


Figure 7.16: Trait physically active plot for Professor-Researcher Scenario.

Figure 7.16 shows the trait physically-active plot for professor-researcher scenario. The trait is associated with more dominant and more pleasant temperament characteristics. The physically active trait is almost 0 in the first 10*seconds* of the interaction. However, it increases tremendously in the later half of the interaction. Since the subject seems quite confident and shows rapid body movements during the interaction, the system finding is quite accurate.

Figure 7.17 shows the trait nurturance plot for professor-researcher scenario. Nurturance is related to giving sympathy, helping others in need, caring for children, and so on. Nurturance is dependent on positive facial expressions display which in turn means positive pleasure. Therefore, during the first 10*seconds* of the interaction, nurturance score is in negative because of negative pleasure value. However, trait value increases in the later half of the duration. The validity of this trait is highly subjective and depends on the context and the scenario. The professor-researcher scenario is a formal scenario with limited context; therefore, the validity of this trait demands a much more diverse scenario.

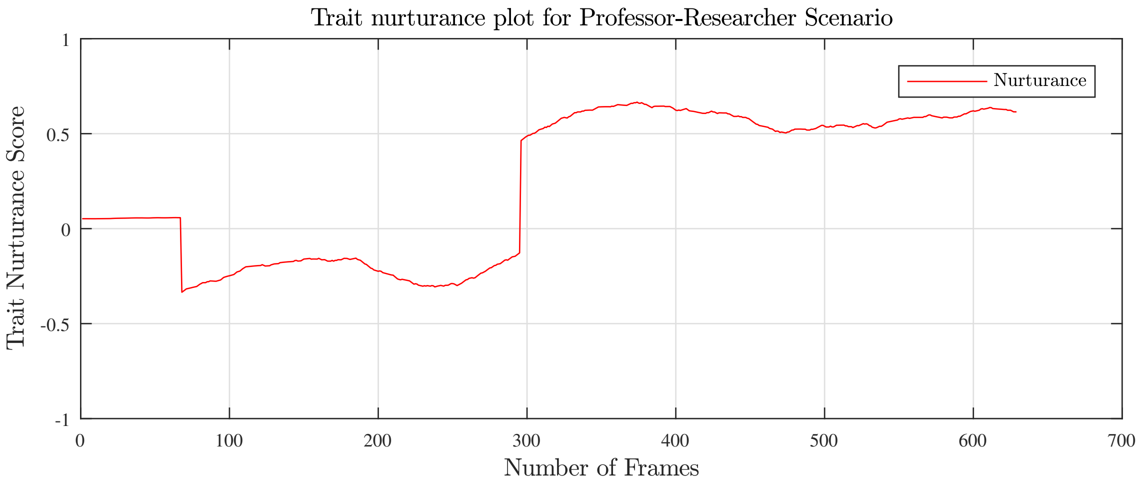


Figure 7.17: Trait nurturance plot for Professor-Researcher Scenario.

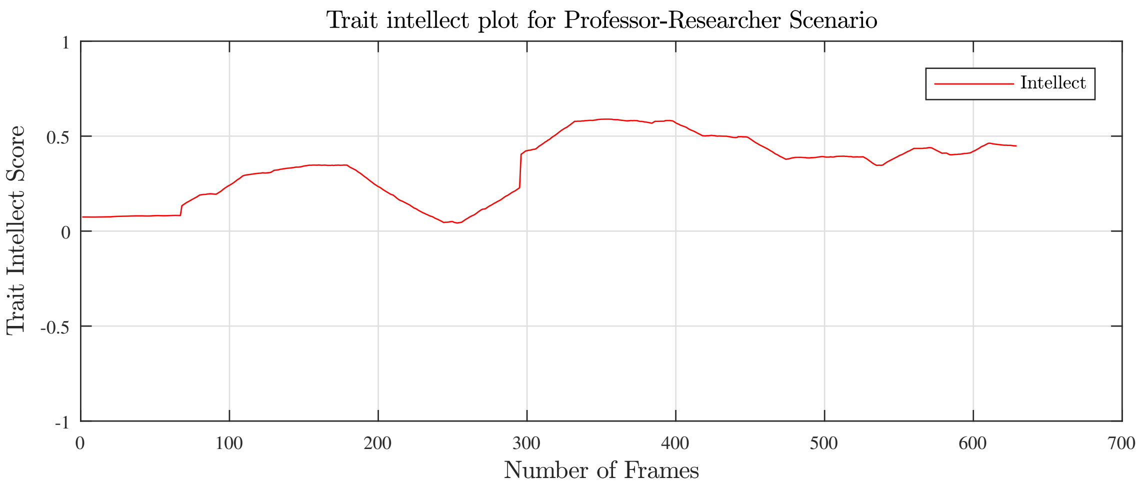


Figure 7.18: Trait achievement plot for Professor-Researcher Scenario.

Figure 7.18 and Figure 7.19 show the trait intellect plot and trait achievement plot for professor-researcher scenario, respectively. Both these traits are strongly correlated with the dominance dimension. Intellect trait defines a person who involves in critical thinking and research with others. Intellects also proposes logical solutions of the problems. As can be seen from Figure 7.18, the intellect trait for the subject is fairly positive almost throughout the duration of interaction. In the context of this scenario, it means that the subject is giving sound reasons and valid excuses for not completing the task.

The similar fact can also be validated from Figure 7.19. Something done successfully with effort, skill, or courage shows achievement. The trait achievement is dependent on dominance. As can be seen from the Figure 7.19, the trait score is positive throughout the interaction. It shows that the subject is able to defend herself successfully which seems true as well.

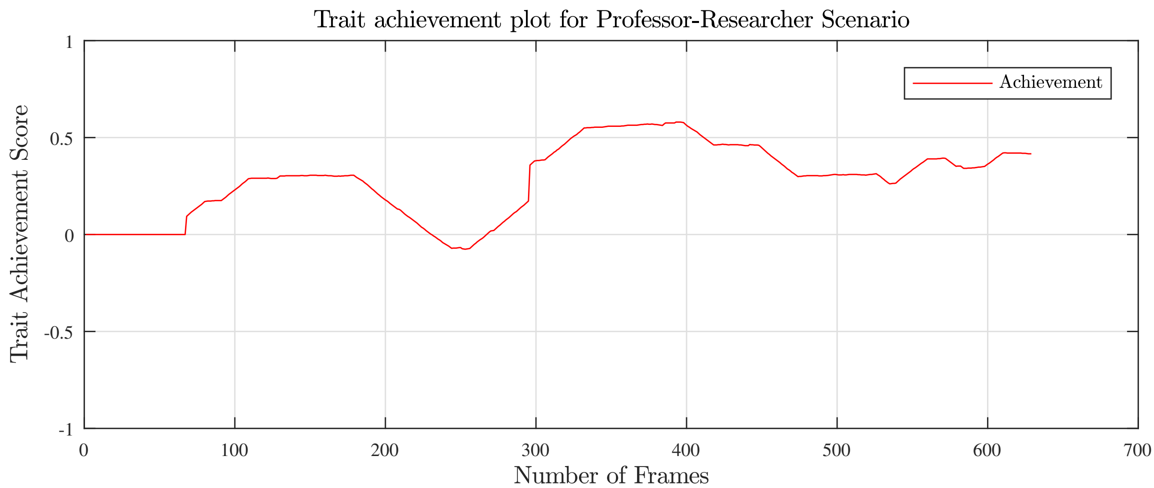


Figure 7.19: Trait intellect plot for Professor-Researcher Scenario.

7.3.4 Performance Evaluation

In order to evaluate the personality assessment system, summative evaluations are used from 5 psychology students. All the interactions are video recorded, and these videos are used for further analysis. After the experimentation, each evaluator is presented with a list of 12 personality traits and their corresponding description as follows:

1. Intellect: A person who engages in critical thinking, research, and reflection about society, proposes solutions for its normative problems, and gains authority as a public figure.
2. Achievement: Something done with effort, skill, or courage.
3. Extroversion: Extroverts are behaviourally more dominant in face-to-face interactions with others.
4. Social Desirability: To answer questions in a manner that is viewed to be favourable by others.

5. Arousal Seeking: A person that looks for excitement, change, new environments, taking the risk, etc.

6. Aggression: Readiness to attack or confront someone.

7. Dominance: Showing power and influence over others.

8. Physically Active: A person that is continuously active, working, organising activities, etc.

9. Anxiety: A feeling of worry, nervousness, or uneasiness about something with an uncertain outcome.

10. Shyness: Nervous or timid in the company of other people.

11. Sensitivity to Rejection: People who are affected easily by the negative remarks of others.

12. Nurturance: Emotional and physical care given to someone.

Observer evaluations have been used to assess human personality traits. Evaluators assess the personality traits of all the subjects using the recorded videos and descriptions of personality traits to establish the ground truth. They use the provided descriptions and their prior knowledge of human behaviour to form an informed judgment about the presence (active) or absence (inactive) of each personality trait. In order to combine the outcomes from each evaluator to generate ground truth, the maximum occurrence of the outcome is used as the final outcome. For example, if two evaluators report the subject's anxiety trait as active and three evaluators report it as inactive, then the ground truth for anxiety trait of the subject is described as inactive. These assessments are compared with the system results. However, the proposed system reports the trait values between -1 to $+1$ range. For evaluation and validation, the trait with positive value is considered as active, and the trait with a negative value as inactive. Table 7.5 shows the recognition rates of each personality trait.

It can be seen from Table 7.5 that extroversion, arousal seeking, trait dominance, physically active, and anxiety traits have higher recognition rates. It is because these traits are distinct and easily distinguishable visually. For example, the system detects the presence or absence of an extroversion trait with 90% accuracy. Most of the subjects that are identified as high on extroversion trait are expressive and physically active with an open body stance during interactions, which in turn yields higher scores on P.A.D. scale. Therefore, the system estimates the extroversion trait as active for these subjects.

On the other side, some of the subjects appear to be nervous and shy during interactions, which results in lower values for P.A.D. scale. Therefore, the system estimates the extroversion trait as inactive for these subjects. Similarly, some subjects show anxiety during interactions by exhibiting dejected and self-touching postures. Subjects are also found to be passive and restless (leaning side to side) during interactions. It has been found after analysis that dominance scores for such subjects have been negative. However, their arousal score is positive due to their movement attributed to restlessness. Therefore, the system estimates the anxiety trait as active for these subjects.

Table 7.5: Personality traits and recognition rates

ID	Trait Dimension	Accuracy (%)
1	Intellect	83.11
2	Achievement	78.22
3	Extroversion	91.11
4	Social Desirability	80.44
5	Arousal Seeking	87.55
6	Aggression	84.88
7	Trait Dominance	89.77
8	Physically Active	92.88
9	Anxiety	90.66
10	Shyness	78.66
11	Sensitivity to Rejection	73.33
12	Nurturance	77.77
	Average	**84.03**

The reason for the wrong assessment of some personality traits lies in the inaccurate recognition of facial expressions. The facial expression recognition (FER) system works accurately when a person is expressing the emotions clearly. During the interaction, the facial expressions of a subject are sometimes wrongly interpreted, which affects the pleasure value. Nurturance and social desirability strongly correlate with pleasure scale as can be seen from equation 3.1, but due to the technical limitation of FER system, these traits achieve low accuracy as depicted in Table 7.5.

Due to the subjective nature of the task, evaluators themselves find it difficult sometimes to have a mutual consensus on these traits. Because of the availability of additional information such as verbal, situational, and contextual cues, evaluators labelled the ground truths of personality traits for each subject accordingly. However, the personality traits system uses only visual information to analyse human behaviour and, therefore, sometimes wrongly reports a subject on a particular trait. For example, sensitivity to rejection trait is highly subjective and needs a context as well to recognise it accurately. Some subjects that exhibit submissive body postures, such as crossed arms and thinking postures, and avoid eye contact during interactions are detected as high on sensitivity to rejection trait, which may not be true in every case. For example, introverts also show similar body postures and head gestures, however, they may not be sensitive to rejection.

Since contextual and situational cues are not considered in this work, the system is not able to differentiate between fake personality and genuine personality. The system assesses personality based on nonverbal cues, which can also be sometimes expressed artificially. Although the subjects are instructed to enact genuinely, some may have faked their responses. Therefore, the personality assessed by the system may differ from the actual personality of the subject, which shows the importance of contextual and situational cues. However, the highly engaging nature of scenarios along with the robot's human-like gestures and expressions, challenge the subjects to respond with minimal artificiality in this work. Hence, the system can assess different personality traits with 84% accuracy.

7.3.5 Robot Adaptivity

The ability of humans to adapt according to the behaviour of their interlocutor has been proven essential for an effective conversation in HHI. For social robots to interact naturally with humans, they must be well-adapted to human behaviour and reactions. The objective is the adaptivity of a social robot based on the interaction partner information. This information may include an interlocutor's profile, emotions, behaviour, personality, and past interactions. Using this knowledge, a social robot should adapt its behaviour accordingly [Beer 14]. An adaptive social robot is expected to have following adaptation capabilities in HRI: understand and show emotions, communication with high-level dialogue, learn/adapt according to user responses, establishing a social relationship, react according to different social situations and have varying social characteristics and roles [Fong 03].

The objective of the thesis in this study from a robot's point of view is to assess different personality traits of humans in real-time and to adapt according to the perceived personality. In this thesis, the personality information of a human is used as a basis for robot adaptivity. In order to realize the adaptivity of the robot, ROBIN, a scenario has been developed, which is about climate change and global warming. ROBIN expresses its views on climate change and looks concerned. ROBIN adapts its sentences, movements, and facial expressions according to the perceived behaviour. Table 7.6 shows the questions and responses of ROBIN during an interaction, while Appendix D, Dialogue D.2, shows the flow of the dialogue.

Table 7.6: Dialogues of ROBIN during weather scenario.

ID	Dialogues of ROBIN during Interaction
1	Have you noticed the change in the weather lately? It is so bizarre. I think after all, global warming is real. Do you like this weather?
2	Why do you think this is (or not) a good weather? Please tell me in detail.
3	Ahh great. I can see that you are excited about the weather. At least someone is happy.
4	You said that you like this weather. However, you don't seem so enthusiastic about it. Maybe something is bothering you. *after a pause* Whatever it is, I hope everything is good.
5	Ahh okay. However, it seems that you are not affected by the weather at all. It is good to see you in high spirits today.
6	Well, I can understand your depressing behaviour. This weather is terrible. Anyway, cheer up now because the weather is going to change next week.
7	Oh no! My battery is running out. I have to rest goodbye

Based on the assessed personality traits, ROBIN changes its sentences, movements, and facial expressions. The scenario begins with an introductory dialogue, as shown in the Table 7.6, dialogues 1 and 2. If a person is found to be high on extroversion and arousal seeking traits, ROBIN uses either dialogue 3 or dialogue 5 of the Table 7.6 depending upon whether the person likes or dislikes the weather. ROBIN also becomes expressive and shows intense smiles and leans towards the person. Similarly, if a person is found to be low on extroversion traits and high on anxiety, ROBIN uses dialogue 4 and dialogue 6, respectively. It also shows comforting and concerned expressions. ROBIN uses dialogue 7 to end the interaction.

8. Conclusion and Outlook

Assessing personality traits is undoubtedly a cognitive aspect that requires intelligence. Recognition of personality traits in the context of Human-Robot Interaction (HRI) is an unsolved topic of research in the field of robotics. Human perception and cognition systems play a motivating role in the visual assessment of personality traits. The system has been developed by bringing together several personality theories and psychology findings. Two different aspects have been considered: *big three* personality traits assessment and subtle personality traits assessment. These aspects have been studied and presented in this thesis.

Human personality traits have been studied by many psychologists extensively, and several theories have been presented in the literature. The proposed system is based on the combination of two psychological theories: *big five* personality theory [McCrae 99] and *temperamental framework* [Mehrabian 96]. Analysis of human behavioural traits requires a variety of visual perceptual skills. These perceptual abilities helps the robot to perceive different human actions. By using a fusion of such different visual perceptual abilities enables a robot to understand human behaviour and personality traits. The thesis is summarized in the following section.

8.1 Summary

This goal of the thesis is to develop a system that enables a robot to synthesize an appropriate behaviour adapted to human personality traits. Human personality is made up of the characteristic patterns of thoughts, feelings, and behaviours that make a person unique. Personality plays a vital role in human-human interaction, and its significance can be better exemplified by two renowned theories from the field of human psychology, i.e., *the chameleon effect* [Chartrand 99] and *the similarity-attraction* theory [Henderson 82].

The chameleon effect explains the non-conscious human tendency to passively mimic the behaviour of one's interaction partner in a social environment. In contrast, the similarity-attraction theory emphasizes that humans are generally attracted to and prefer the company of others who maintain morals and attitudes similar to their own. For example, it is quite often observed that there exists a sense of shared personality among friends than among random pairs of strangers. Similarity-attraction theory can be observed in people with thought processes, such as not feeling alone in their belief, or the ability to predict the future behaviour of similar people in order to access the "window of bias" for enhanced relationships and validation of attraction. Subsequently, people tend to change their behaviour according to their interlocutor's behaviour. If he/she is talkative and expressive, one also tends to be more expressive.

There is another compelling theory called complementary attraction which describes that individuals also attract towards those people whose personalities are complementary to their own personalities [Isbister 00]. The good example of complementary theory is the long-term relationships between two persons with particular roles such as in marriages and in office environment. Similarly, several studies have also found the relationship between human personality and the proximity behaviour. Tapus et al. [Tapus 08] have found that people with extroversion personality type are more lenient with their personal space invasion by a robot than introverts.

Hence, the assessment of human personality traits is highly critical for a robot to adapt its behaviour appropriately during human-robot interaction. In order to build such a system, a personality assessment architecture has been modelled using psychology and cognitive studies. This architecture is also responsible to enable a robot to adapt its behaviour and behave in an appropriate manner.

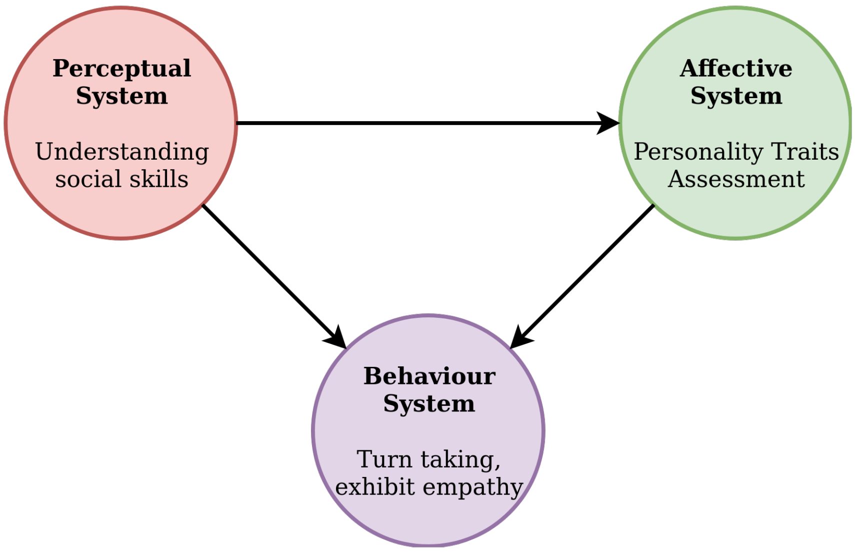


Figure 8.1: Perceptual System helps the robot to understand social skills, **Affective System** helps the robot to recognize personality traits, and **Behaviour System** helps the robot to behave intelligently using the information from perceptual and affective system.

The implemented personality trait assessment system is organized in three separate levels, as illustrated in Figure 8.1. The first level, known as perceptual level, is responsible for enabling the robot to perceive, recognize and understand human actions in the surrounding environment in order to make sense of the situation (see Chapter 5). The second level, known as affective level, helps the robot to connect the knowledge acquired in the first level to make higher order evaluations such as assessment of human personality traits (see Chapter 6). By using the information from the perceptual and affective level, the last level, known as behaviour level, enables the robot to synthesize an appropriate behaviour adapted to human personality traits.

Visual Perceptual System

Analysis of human behavioural traits requires a variety of perceptive skills. These perceptive skills, which include understanding of human nonverbal cues such as expressions, gestures, postures and many more, over time help the robot to perceive different human actions. Usually, humans detect and recognise different behavioural traits of their counterparts by analysing their nonverbal cues over time. Extrovert people tend to be quite active, and their excessive hand movements during conversations most often show confidence and control [Oberzaucher 08]. Similarly, other *percepts* such as body posture, hand gesture, and facial expressions, play an important role in extracting the emotional state of an interlocutor.

Human Postures

Posture plays an important role in expressing human personality traits. Some behavioural cues can be easily recognized from postures. For example, upright posture, open arms, and a genuine smile convey ease and confidence, which is the sign of extroversion [Kuhnke 12]. Similarly, crossed arms posture shows that the interlocutor is reserved and is trying to block himself from opening to other people. Posture recognition uses RGB-D sensor to extract human joint positions using the depth stream. From feature extraction perspective, using depth data to extract features is more useful than colour images. The reason to use depth information is that the colour images are always hampered by lighting conditions, person's clothes and partial occlusions of human body. A significant contribution of this thesis is a novel feature extraction method that converts human joint positions into joint angles in order to make the system invariant to height, scale, position and physique of the person. Comparing with state-of-the-art algorithms, the proposed posture recognition has shown robustness and efficiency. Experiments conducted in real-time have shown very high accuracy better than the most state-of-the-art systems. Additionally, the experiments have shown that the proposed approach is quite robust, which takes only $0.25ms$ on average to estimate the posture. Furthermore, the overall accuracy of the developed approach is more than 95%.

Proximity, Body Movements and Speech Duration

The distance that the interaction partner maintains during their conversations conveys important messages to both of the interaction partners and other people. Mehrabian [Mehrabian 69] has stated that the distance between an individual and his addressee is a decreasing linear function of the degree of liking of the addressee. Moreover, individuals high on extroversion prefer to sit or stand close to the conversation partner [Knapp 13] [Hargie 16]. Using the depth data of tracked human, the proximity of a person relative to the robot has been estimated. It has been shown from the experiments that the estimation of proximity is highly efficient and robust. Additionally, the system takes only $0.005ms$ on average to estimate this *percept*.

Similarly, body movements during an interaction is an important *percept* which sheds light on the spirit of a person. According to Nass et al. [Nass 01], people, when aroused, show frequent body movements. Moreover, extroversion is related to more frequent and more rapid body movements [Oberzaucher 08]. Analysing upper skeletal joint angles over time provides the information of body movements. This percept also uses depth information to analyse human movements. Therefore, these features are illumination, scale, and position invariant. Furthermore, this *percept* is highly efficient and robust as it takes only 0.012ms to compute.

According to [Argyle 13], extroverts talk more and they produce more words and talk longer when they have the turn. Extroverts talk faster, louder, with shorter pauses and with a higher pitch [La-France 04] [Matsumoto 13]. This feature can easily be correlated with the duration of the speech. A person that talks for long duration are known to be an extrovert and a person who talks less is termed as an introvert and a loner, who avoids people. Due to the absence of speech recognition, this *percept* has been manually estimated (see Chapter 5). In the personality assessment experiments, this *percept* has been found to be quite significant.

Head Poses and Head Gestures

This *percept* plays a huge role in the assessment of personality traits. For example, self-centred people convey a sense of superiority by lifting their head and tilt it backwards, and people perceive them as haughty. Similarly, they also raise their head and thrust their chin forward [Kuhnke 12]. Moreover, agreeable persons nod a lot during the conversation in order to encourage the interlocutor to continue, while neuroticism and introversion traits are positive correlated to looking down and avoiding mutual eye gaze [Knapp 13]. The percept has been extensively used for the assessment of personality traits.

Facial Expressions

According to [Ruch 94], extroversion is associated with more frequent and more intense smiles [Argyle 13]. Neuroticism & Introversion are related to low expressiveness [Argyle 13]. Agreeableness is positively correlated with laughter and a sympathetic facial display [La-France 04]. This *percept* plays an important role in assessing the internal emotional state and personality of humans. We use expressions to convey our thoughts and feelings during interpersonal communication. This *percept* has been used in the assessment of personality traits and it has shown to be a significant contributor in perceiving 'trait pleasure' from human face.

Affective System

The affective system of the robot is responsible for analysing human personality traits. To the best of our knowledge, this thesis is the first work in the field of human-robot interaction that presents an automatic assessment of human personality traits in real-time using visual information. Using psychology and cognitive studies, many theories has been studied. Two theories have been been used to build the personality trait assessment system: *Big Five* personality traits assessment and temperament framework for personality traits assessment.

BF Personality Traits Assessment

Different psychology theories of personality traits have been studied, and ultimately, the big five personality traits theory has been considered [McCrae 99]. As mentioned earlier, BF personality theory defines a person in five unique dimensions in which each dimension is a continuum. The research goal is to enable a humanoid robot to analyse human personality traits and adapts its behaviour.

In order to understand the role and significance of nonverbal cues, a psychology survey on human behaviour analysis has been done. According to various psychologists, several human nonverbal cues indicate the emotional state as well as the personality traits of a person. A dataset has been generated using 15 subjects. Each nonverbal cue is correlated using Pearson's correlation coefficient with the traits to analyse their relationship between them and to validate the psychology claims. It has been found that almost all of the psychology facts are validated. Some percepts do not play any role towards personality traits assessment such as, disgust. These percepts are discarded before the classification task. To validate the system, multiple scenarios has been generated. Subjects have been asked to act in the scenarios. Experimental results have shown that our system can recognize BF personality traits with more than 90% accuracy.

Subtle Personality Traits Assessment

As noted earlier, big five personality traits are generalised traits of a person. However, these traits do not distinguish a subset of big five categories. For example, aggression, dominance and physically active are sub-traits of self-centred and extroversion dimension. The assessment of these subtle traits is highly significant and relevant in the context of HRI. This thesis have proposed to use the P.A.D. emotional space for the assessment of human personality traits using the framework presented in the literature [Mehrabian 96].

Using the three dimensions, pleasure, arousal, and dominance, as explained in Chapter 3, the author has formulated 59 individual measures that correspond to human personality traits. It has been demonstrated that traits are symmetrically related to one another based upon the P.A.D. dimensions. Although the formulated traits are of a wide range, only 12 out of 59 traits are realised in this thesis. These traits are chosen according to the experimental restrictions and based on the knowledge of nonverbal cues associated with them.

To use the temperament framework, human emotional state is represented in pleasure, arousal and dominance emotional space. Using psychology and cognitive science studies as the reference, perceptual skills of the robot have been used to estimate the P.A.D. dimensions. The regions in the P.A.D. emotional space represent human personality traits. Based on these regions, different personality traits of a person are computed. As in BF personality traits validation, multiple scenarios have been generated to validate the performance of the temperament framework. Overall, the system is able to achieve 84% accuracy for 12 subtle personality traits.

Behaviour System

The robot behaviour system is responsible for the physical adaptivity of the robot. There are many ways in which the robot can express its behaviour such as, gestures, postures, expressions, gaze and speech. Using personality theories from psychology, the robot can adapt appropriately, as can be seen from experiments (see Chapter 7). As mentioned earlier, robot uses the similarity-attraction principle and behaves with similar personality type. For example, if the person is found out to be extrovert, the robot also behaves like an extrovert. However, it also uses the complementary attraction theory to adapt its behaviour and complement the personality of the interaction partner. For example, if the person is found out to be self-centred, the robot behaves like an agreeable in order to flourish human-robot interaction.

8.2 Future Work

The work presented in this thesis can be further extended in many directions. These extensions are discussed in the following paragraphs.

Extension of Personality Assessment System

The assessed personality traits in this thesis are of wide variety. These traits are used extensively in human-human interactions. However, one important factor that has been ignored in this thesis is the paralanguage cues. These cues also represent human personality traits and if added with the existing system, these cues will enhance the accuracy of the personality trait assessment system.

Moreover, speech also play huge role towards personality. Having a speech recognition system can help the robot to understand the context more easily. Therefore, this is one other aspect where personality trait system can be improved.

Context Aware Perception

Currently, the robot interacts based on what it perceives. It does not have a memory to store its experiences, important information and interaction outcomes. Creating a memory for a robot can help the robot to personalize its behaviour according to the identity of the person. Robot can store information about the interests of person, his/her personality, his/her academic background and so on. This information can help the robot to talk to each person in a more personalized way.

Emotion Based Control Architecture

Although it is not a new concept, however, building such an architecture helps the robot to synthesize its behaviour automatically using emotions and external stimulus. This can help the robot to be adaptable and flexible. It can adapt to varying environmental conditions. It can be done by changing parameters to change the behaviour of the components and the sub-systems.

In conclusion, this thesis contributes to the development of personality trait assessment system for a socially interactive robot. Transferring of psychological and cognitive personality models and theories to a robotic system enables the robot to behave more naturally and appropriately. Enabling the robot with personality trait assessment system is the first step towards intelligent human-robot interaction.

A. The Humanoid Robot, ROBIN

To realise human-robot interaction, the humanoid robot, ROBIN, of the robotics research lab, TU Kaiserslautern has been used, as shown in Figure A.1. ROBIN is a human-size upper body robot. It is equipped with a backlit projected face that can express more than 6 basic facial expressions. The arms make use of pneumatic muscles to show fluid motion and express human-like gestures.

ROBIN also has intelligent human-like hands with 8 degrees of freedom (DoF) in fingers for flexion, the finger spread, thumb pitch, and thumb roll. There are 2 DoF in the wrist (pitch and roll), 1 DoF in elbow (pitch) and 3 DoF in the shoulder (pitch, roll, and yaw). ROBIN's compliant neck and torso both have 3 DoF. Pneumatic muscles in the arms are powered by external air supply. For the perception task, an RGB-D sensor, Asus Xtion Pro, is installed on the chest of ROBIN. ROBIN also has an onboard pc that is responsible for control movements, expressions, speech and some perception algorithms. A stand-alone Intel Core i7 running at 3.40GHz with 16GB RAM has been used to process the RGB-D data.

A.1 Interface

The interface of ROBIN is build on tritium architecture. The architecture has 3 major components: control communication manager, communication hub and IOserve. Control function manager access the control scripts and the robot definitions, IOserve is responsible for the low-level hardware control (CAN bus etc.) and communication hub is responsible for the communication between web-based graphical interface and the robotic system. Figure A.2 shows the tritium architecture of the robot.

ROBIN has an interface that allows the FINROC framework on the external computer to control the robot. There exists two nodes in robot's internal software architecture, which can receive commands from external applications via HTTP requests. The first node is a *control* node which is used to control the robot. It is responsible for controlling joint angles, expressions, face guises, gaze, cheek's colour, facial action units and blinking rate. Moreover, it is also responsible for text-to-speech feature.

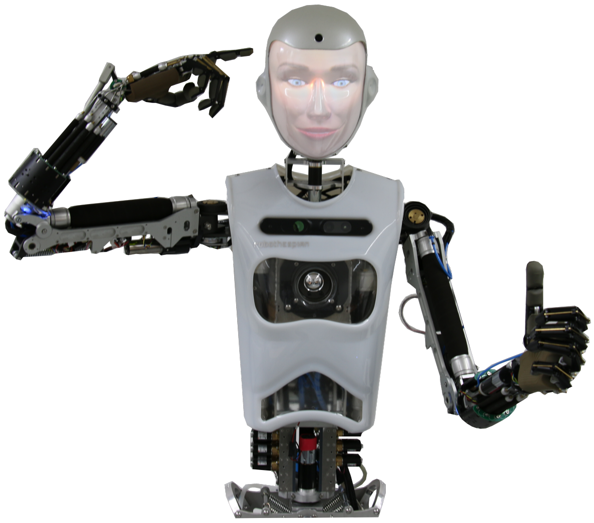


Figure A.1: The interactive humanoid robot, ROBIN, of TU Kaiserslautern.

On the other hand, the second node is called *sense* node. It is used to sense and read the current encoder sensor values. This node is also responsible to sense the environment using the HD camera installed on the head of the robot. The *sense* node captures the images of the environment and passes it to the robot's internal perception system.

A.2 Internal Perception System

From the *sense* node, ROBIN internal perception system receives colour stream. The system uses SHORE library to extract valuable information from human faces. SHORE is a face detection software which allows for the quick detection of faces. It can estimate gender, age and facial expressions in real-time. It can also detect facial features such as eyes and mouth and their state (closed/open). It can estimate four facial expressions: happy, sad, surprised and angry. The outcome of the SHORE library is send to the FINROC framework via an HTTP requests.

A.3 Simulation

ROBIN is provided with a web-simulation interface. The interface allows a developer to develop several behaviours of the robot. It provides access to manoeuvrer facial expressions and action units, body joints and the speech. There is also a facility to use pre-stored postures, limbs movements, and sequences, which can be directly used and modified according to the need. The verbal behaviour can be invoked either by text-to-speech plug-in in several accents or by using audio files in .wav format. This interface provides an easier method to generate content remotely. This content can be uploaded to the robot via robot's web-interface. Figure A.3 shows the simulation environment.

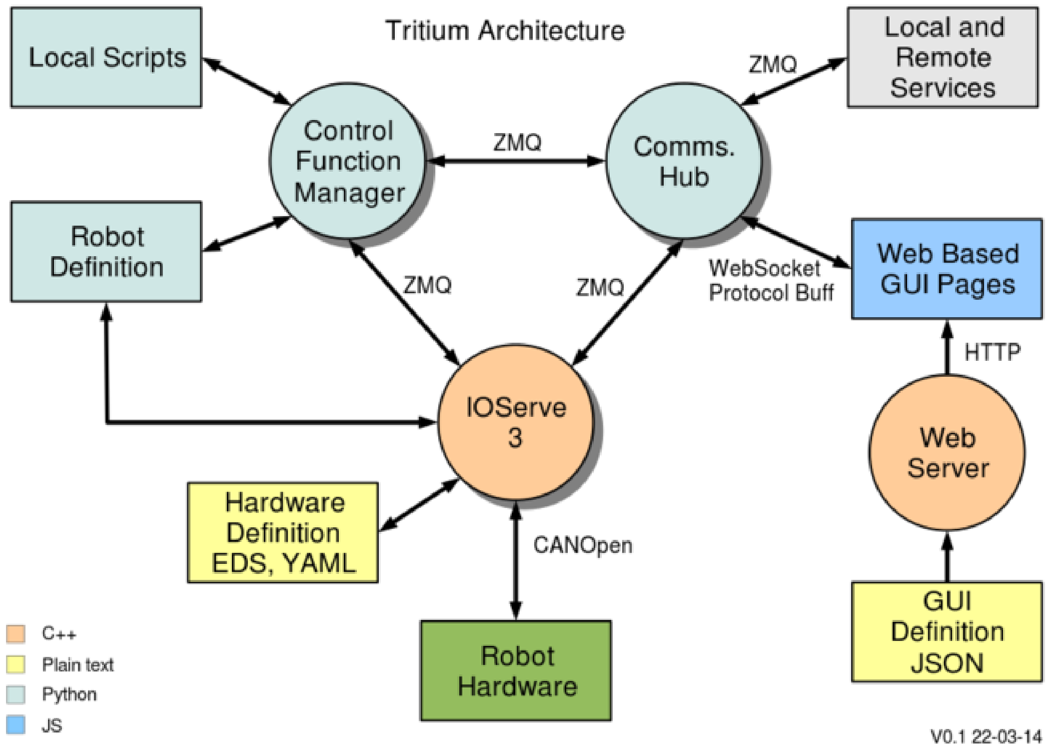


Figure A.2: Tritium architecture of the humanoid robot, ROBIN

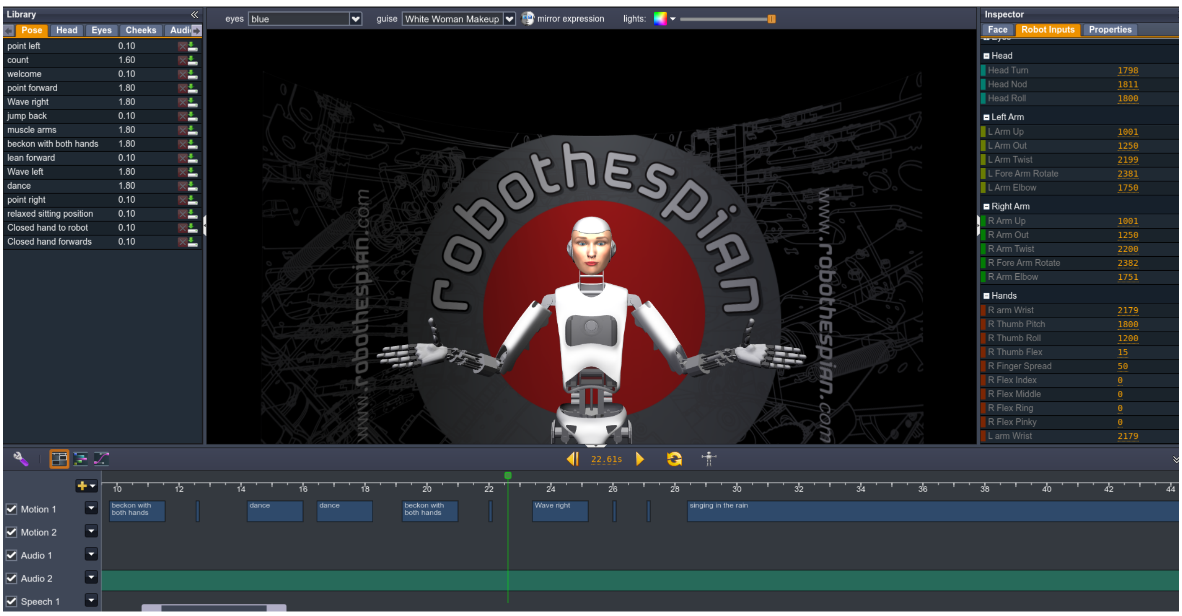


Figure A.3: Simulation environment of virtual robot, ROBIN, for generating content.

B. Robot Framework, FINROC

The personality trait assessment system reported in this thesis has been developed in the FINROC robotic framework. FINROC is a modular C++/Java framework for robot control systems. FINROC has been developed at Robotics Research Lab, TU Kaiserslautern. The structural elements of FINROC applications are *modules* (basic components), *groups* (composite components) and *parts* (executable processes) [Reichardt 13] [Reichardt 17].

The basic FINROC applications are programmed in *modules*. A project can have multiple applications that are responsible for several different tasks. These separate *modules* are placed in *groups* to maintain a clear application structure. These *modules* and *groups* are interconnected with a set of *ports*, input ports and output ports. The input ports can be connected to output ports if they have same data type. FINROC framework also differentiates between sensor and controller data. In this way, it supports clear visualization and structure of the application.

FINROC provides two graphical tools for visualization: *finstruct* and *fingui*. *Finstruct* is a helpful graphical interface which enables the developer to connect, instantiate and remove components of the application at runtime. Apart from visualization, these advantages make *finstruct* a handy debugging tool. It allows the developer to visualize the incoming data from sensor ports, sending control signals and activating/deactivating several components. Figure B.1 the overview of the whole perception system of the robot. Figure B.2 shows the low-level perception system of ROBIN. Appraisal system of the robot is displayed in Figure B.3. The overall architecture of the ROBIN is shown in Figure B.4.

Fingui tool is a sophisticated end user graphical user interface which helps the user to visualize sensor data. There exists several widgets that can be used to visualize data, such as video renderer is used to visualize incoming video stream from a camera. Several widgets can also be used to transmit control signals to application. Figure B.5 shows the graphical interface of robot's control system. Figure B.6 shows the graphical interface of the robot's perception system and Figure B.7 shows the graphical interface to visualize scores of subtle personality traits.

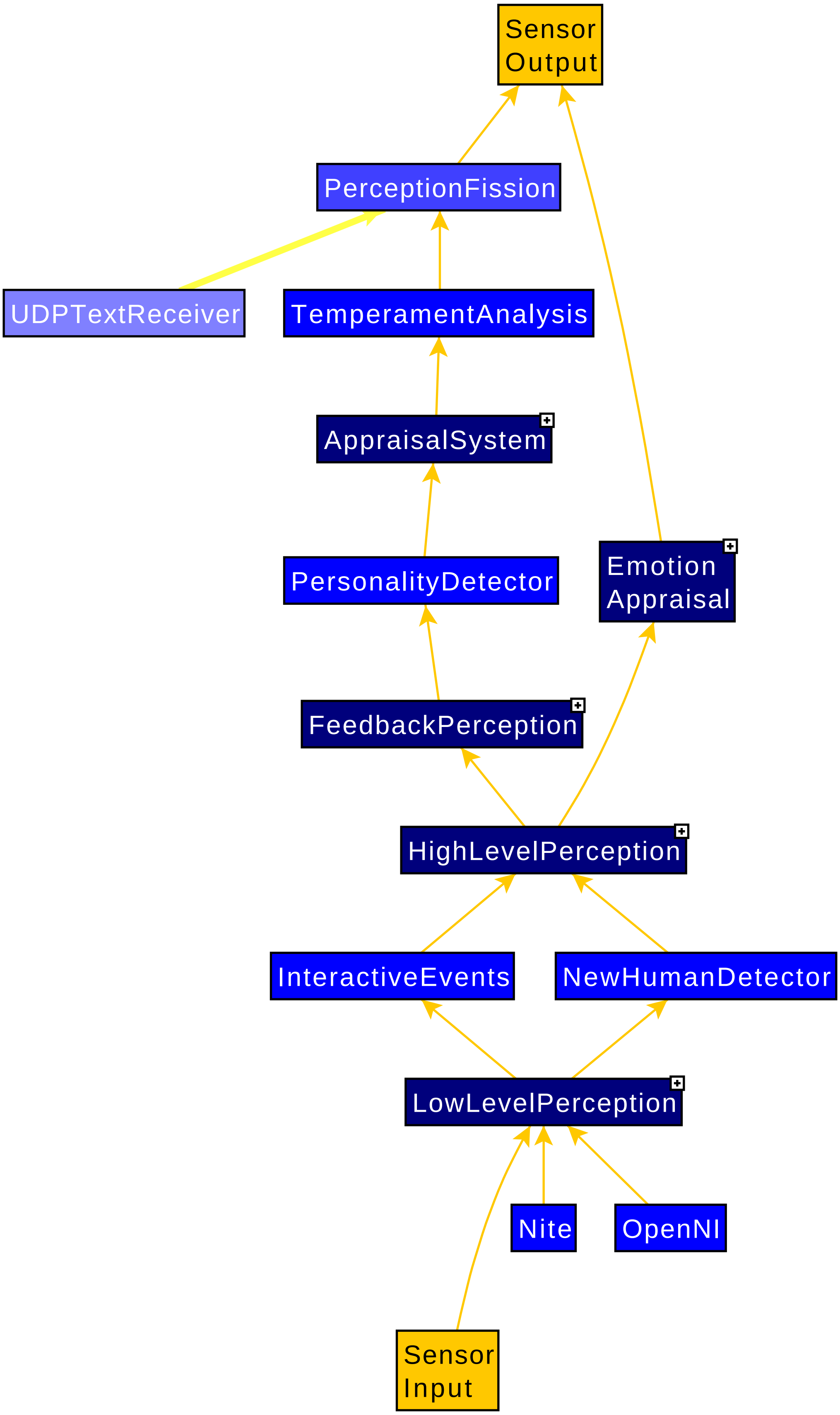


Figure B.1: The finstruct showing the perception system of ROBIN.

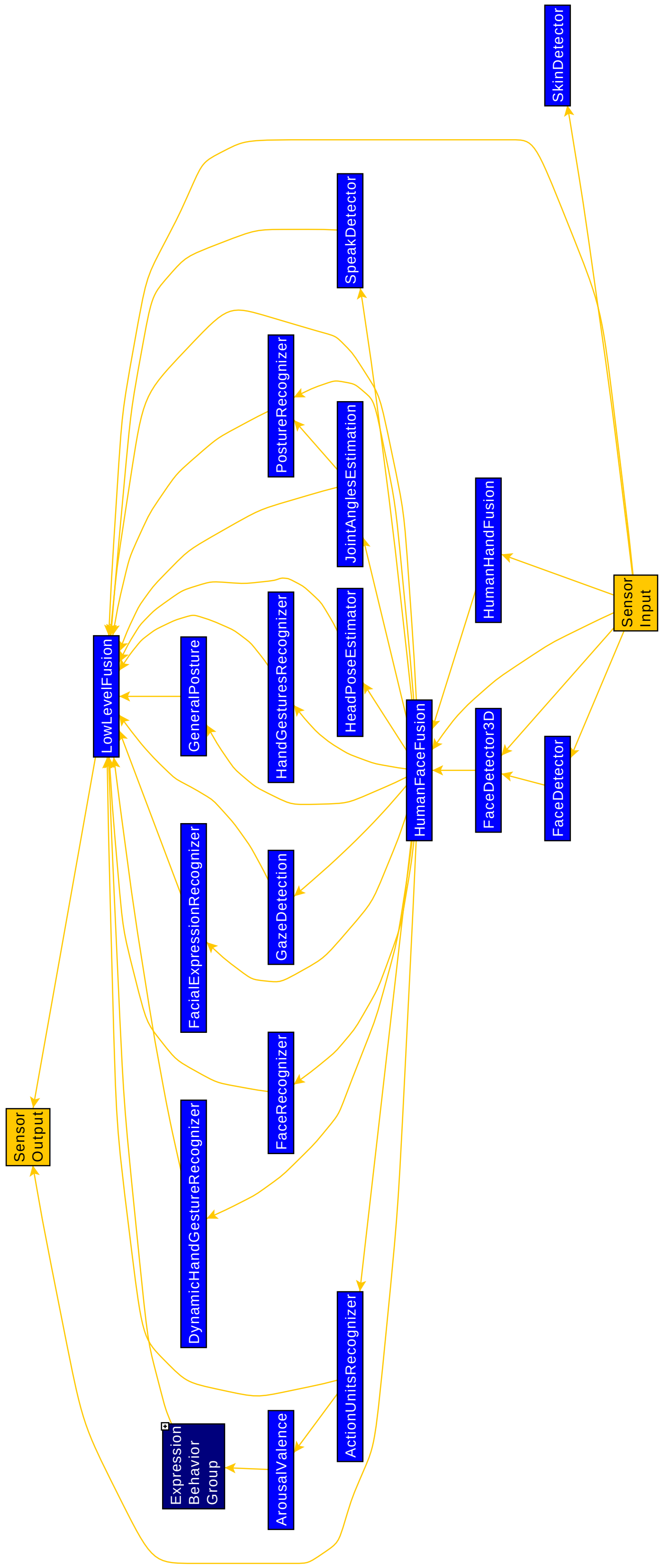


Figure B.2: The finstruct showing the low-level perception system of ROBIN.

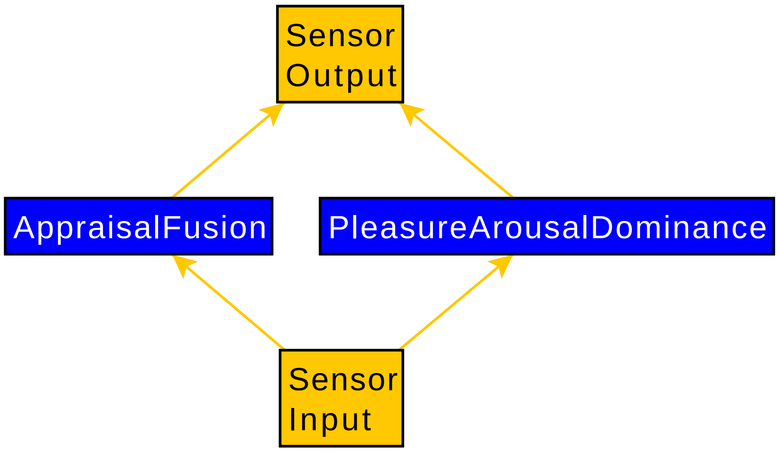


Figure B.3: The finstruct showing the appriasal system of ROBIN.

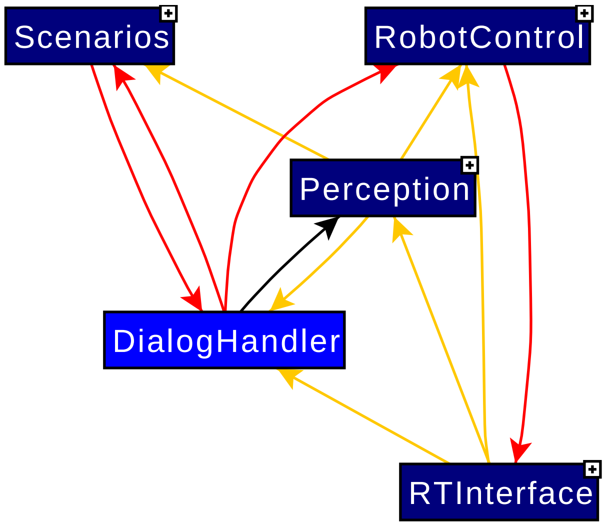


Figure B.4: The finstruct showing the overview of the ROBIN system.

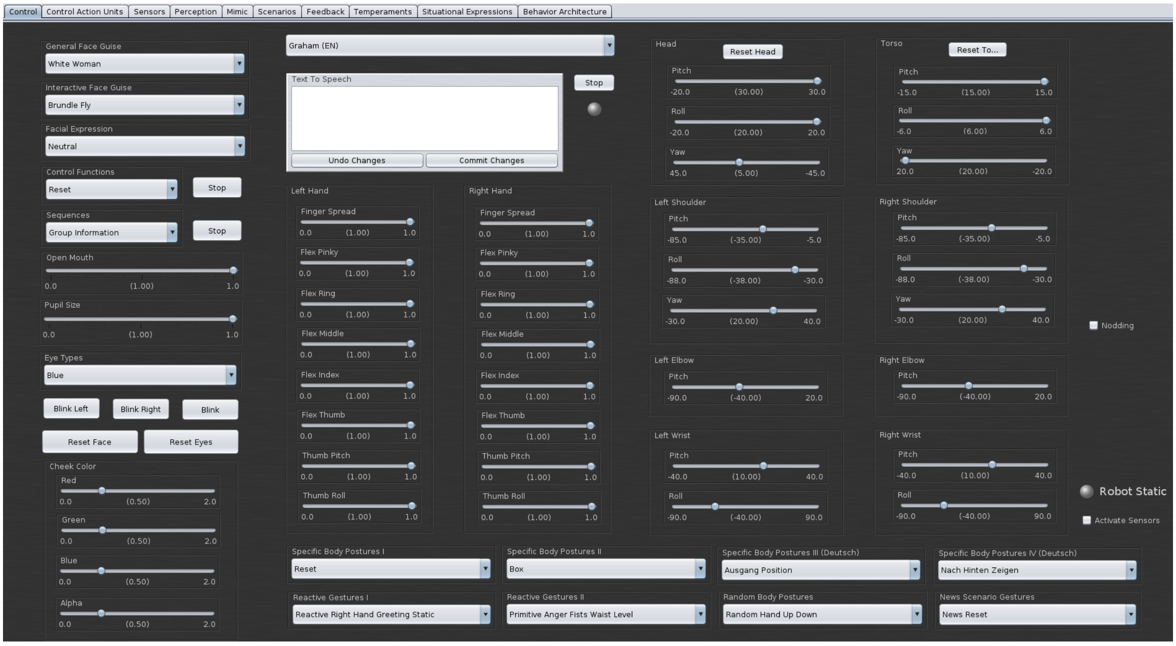


Figure B.5: The fingui showing the control interface of ROBIN.

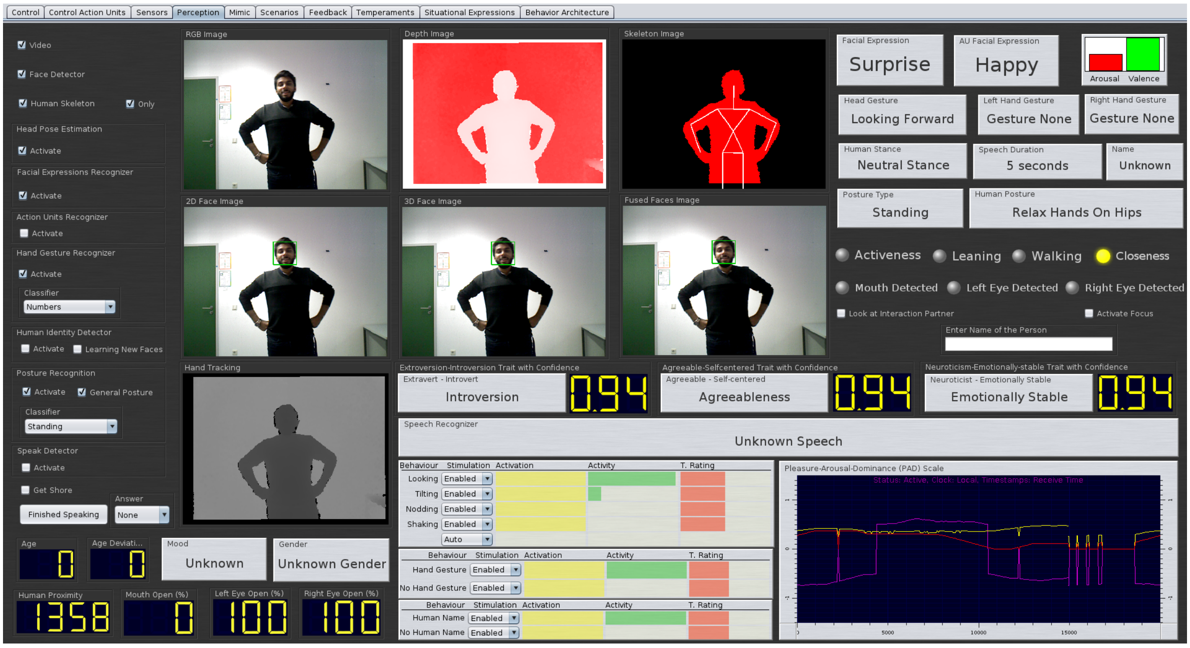


Figure B.6: The fingui showing the perception interface of ROBIN system.

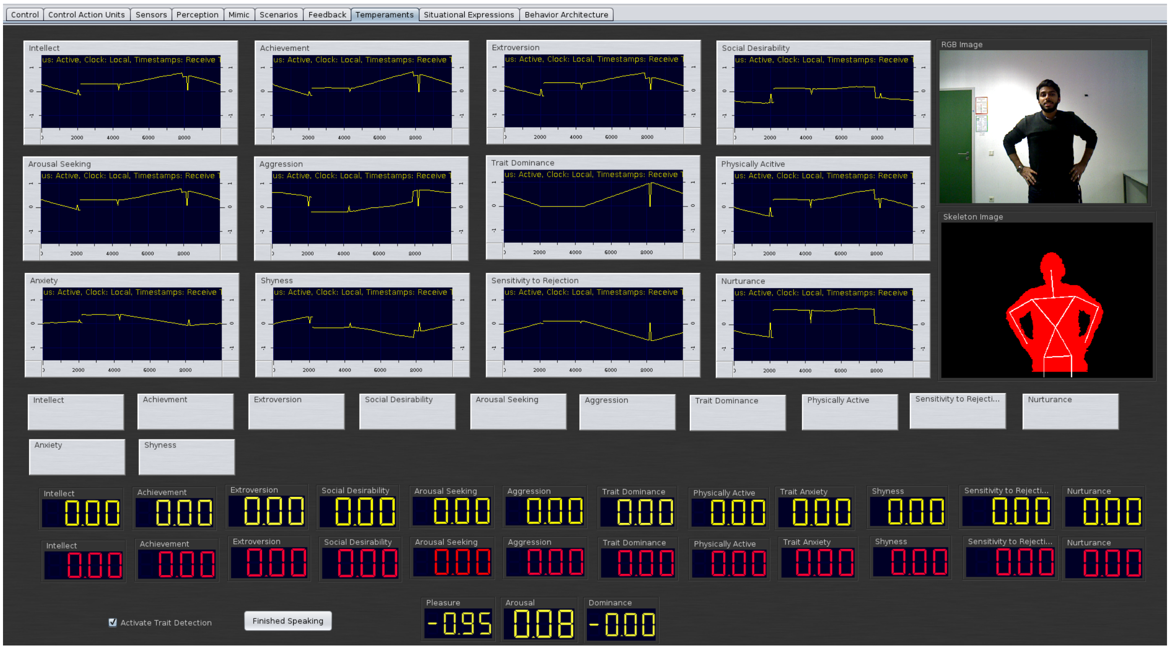


Figure B.7: The fingui showing the visualization of subtle personality traits scores.

C. Perception Softwares

C.1 OpenNI Library

Open Natural Interaction (OpenNI) is an open-source software that interacts with various RGB-D sensors to retrieve visual information[1]. The information usually consists of a colour image, and a depth image. The depth image pixel is represented by a 16 bit value, and each channel for the colour image is represented by 8-bits. The depth value usually represents in millimetres. Figure C.1 shows the colour and depth stream of the scene.

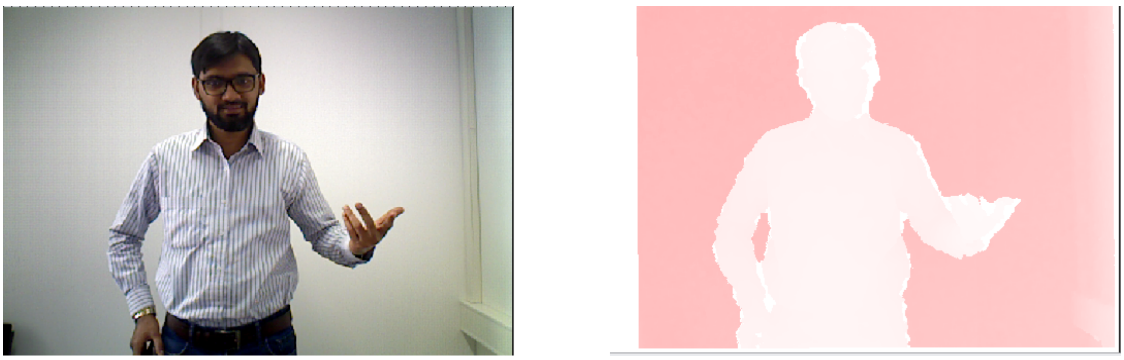


Figure C.1: Colour and depth stream of the scene from OpenNI software.

C.2 NiTE Library

Natural Interaction Middleware (NiTE) version 2.2 is a middleware library pack that uses OpenNI to track human skeleton joint positions. The primary purpose of the User Tracker algorithm is to find all of the active users in a specific scene. It individually tracks each human it finds, and provides the means to separate their outline from each other and from the background. Once the scene has been segmented, the User Tracker is also used to initiate Skeleton Tracking and Pose Detection algorithms.

[1] https://structure.io/openni

Each user is provided an ID as they are detected. The user ID remains constant as long as the user remains in the frame. If a user leaves the field of view of the camera, or tracking of that user is otherwise lost, the user may have a different ID when he is detected again. There exists no mechanism that provides persistent recognition of individuals when they are not being actively tracking.

This library can also track the hands and separate users from the background[2]. HandTracker provides access to all algorithms relates to tracking individual hands, as well as detecting gestures in the depthmap.

The core of the hand tracking is an algorithm that finds human hands in each from of the depthmap, and reports the position of those hands in space. This can be used for simple detection of higher level gestures and implementation of gesture based user interfaces. Unlike full body tracking algorithms, handpoint based tracking works on users that are sitting and does not require a full body be visible.

Gesture tracking is generally used to initiate hand tracking. It allows detection of gestures in the raw depth map, without requiring hand points (in contrast to higher-level gestures that might be used to implement a UI using handpoints). These gestures can be located in space to provide a hint to the hand tracking algorithm on where to start tracking.

The output of the HandTracker occurs one frame at a time. For each input depth frame, a hand tracking frame is output with hand positions, gesture positions, etc. The hand gesture recognizer in this thesis uses handTracker to track the hand. The NiTE library returns a vector containing the three-dimensional hand center position. It also tells the number of hands detected and number of persons detected.

This library has been used to extract human joint positions for the posture recognition task, and also used to extract hand center positions for hand segmentation and gesture recognition task. Figure C.2 shows the output of NiTE library in which person is tracked as well as his hands.

Figure C.2: Tracked skeleton and tracked hand using NiTE middleware library.

[2] http://openni.ru/files/nite/

D. Interactive Dialogues

A dynamic dialogue system guides the interaction between humans and robots. This dialogue system has been developed by Koch et al. [Koch 07]. The system reads the dialogue file, which describes the flow of dialogue and builds a finite state machine (FSM). The states of the FSM represent a set of activities that the robot must do. State transitions are triggered either by a sensor input from the perception system or as a result of a callback from the application.

D.1 Professor-Researcher Dialogue

The Professor-Researcher dialogue has been used to evaluate the subtle personality traits of interaction partner. The scenario involves an interaction between ROBIN, role-playing as the professor, and the participant, role-playing as a researcher. Listing D.1 shows the dialogue file for the interaction experiment described in Section 7.3.

The dialogue starts in the line 10 where it checks whether any person is present in the screen. After human detection, the robot greets by waving its right hand. This information is passed in the ‘gesture’ variable, as illustrated in the line 16. Robot starts the scenario related speech in the line 43. After every sentence, a relevant gesture is also passed depending on the context using the ‘gesture’ variable. After robot finishes its talk, it starts analysing person’s personality traits by activating the traits detection module, as illustrated in line 72. It also checks whether the person has finished answering as shown in line 75. Once the robot has detected that the person has stopped answering, it deactivates the trait detection module in order to conserve the system resources, as shown in the line 87. The robot ends the interaction in the line 308 by saying ‘Good Bye’.

Listing D.1: Dialog description used in an interaction experiment described in Section 7.3.

```
1  <?xml version="1.0" encoding="UTF-8"?>
2  <data xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
3      xsi:noNamespaceSchemaLocation="sources/cpp/libraries/dialog_system/etc/
           DoAmi_dialog_description_scheme.xsd">
4
5      <platform type="desktop">
6
7       <dialog dialog_id="boss_employee_scenario">
8           <SUI_dialog>
9
10     <mark name="start" />
```

```
    <if port="humans_exist" value="1" relation=">=">""<goto mark="initiate_interaction" /></if>
    <else>""<goto mark="start" /></else>

<mark name="initiate_interaction" />
    <set_port name="robot_facial_expression" value="0" />
        <set_port name = "gesture" value = "reactive_right_hand_greeting_static" />
    <prompt>Hello Zuhair!</prompt>
    <wait time="500000" />
    <set_port name = "gesture" value = "neutral" />
    <if port="interaction_partner_x" value="1600" relation=">=">""<goto mark="check_the_distance" /></if>
    <else>""<goto mark="intro" /></else>

<mark name="check_the_distance" />
    <set_port name="robot_facial_expression" value="11" />
    <set_port name="look_at_interaction_partner" value="1" />
    <prompt>Come here. We need to talk.</prompt>
    <wait time="3000000" />
    <set_port name="robot_facial_expression" value="0" />
    <if port="interaction_partner_x" value="1600" relation=">=">""<goto mark="wait_for_distance" /></if>
    <else>""<goto mark="intro" /></else>

<mark name="wait_for_distance" />
        <set_port name="robot_facial_expression" value="13" />
    <set_port name = "gesture" value = "come_here_right"/>
    <prompt>Zuhair you need to come near. Let us not make this a public event.</prompt>
    <wait time="5000000" />
    <if port="interaction_partner_x" value="1600" relation=">=">""<goto mark="upset" /></if>
    <else>""<goto mark="intro" /></else>

<mark name="intro" />
    <set_port name="look_at_interaction_partner" value="0" />
    <set_port name="robot_facial_expression" value="1" />
        <set_port name = "input_dialog_string" value = "I gave you a month time with clear directions to finish the task."/>
    <set_port name = "gesture" value = "pointing_front" />
    <wait time="500000" />
    <set_port name = "gesture" value = "%out_random_head_gesture" />
    <prompt>%input_dialog_string </prompt>
    <wait time="500000" />

    <set_port name="robot_facial_expression" value="1" />
    <set_port name = "input_dialog_string" value = "All you had to do was, to spend quality time and effort to successfully complete the task."/>
    <set_port name = "gesture" value = "sign_of_arrogance" />
    <wait time="500000" />
    <set_port name = "gesture" value = "%out_random_head_gesture" />
    <prompt>%input_dialog_string </prompt>
    <wait time="500000" />

    <set_port name="robot_facial_expression" value="10" />
    <set_port name = "input_dialog_string" value = "Why is it so hard to finish on time?"/>
    <set_port name = "gesture" value = "%out_random_hand_gesture" />
    <set_port name = "return_gesture" value = "%out_random_hand_gesture" />
    <wait time="500000" />
    <set_port name = "gesture" value = "%out_random_head_gesture" />
    <prompt>%input_dialog_string </prompt>
    <wait time="500000" />

    <set_port name="robot_facial_expression" value="1" />
    <set_port name="activate_posture" value="1" />
    <set_port name="activate_facial_expressions" value="1" />
    <set_port name="head_pose_estimation" value="1" />
    <set_port name="get_shore" value="0" />
        <set_port name="activate_traits_detection" value="1" />
    <wait time="1000000" />

    <mark name="check_2_answer" />
    <if port="out_finished_speaking" value="1" relation="==">""<goto mark="part_2" /></if>
    <elseif port="robot_static" value="1" relation="==">""<goto mark="perform_gesture_1" /></elseif>
    <else>""<goto mark="check_2_answer" /></else>

<mark name="perform_gesture_1" />
    <set_port name = "gesture" value = "random_open_up_down" />
    <wait time="500000" />
    <set_port name = "gesture" value = "random_head_shaking" />
    <goto mark="check_2_answer" />

<mark name="part_2" />
        <set_port name="activate_traits_detection" value="0" />
    <set_port name="activate_posture" value="0" />
    <set_port name="activate_facial_expressions" value="0" />
    <set_port name="head_pose_estimation" value="0" />
    <set_port name="get_shore" value="0" />

    <set_port name="robot_facial_expression" value="2" />
    <set_port name = "input_dialog_string" value = "But this is unacceptable."/>
    <set_port name = "gesture" value = "%out_random_hand_gesture" />
    <set_port name = "return_gesture" value = "%out_random_hand_gesture" />
    <wait time="500000" />
    <set_port name = "gesture" value = "%out_random_head_gesture" />
    <prompt>%input_dialog_string </prompt>
    <wait time="500000" />

    <set_port name="robot_facial_expression" value="10" />
```

```
    <set_port name = "input_dialog_string" value = "If you had issues, it was your responsibility to convey the issues to your seniors."/>
    <set_port name = "gesture" value = "%out_random_hand_gesture" />
    <set_port name = "return_gesture" value = "%out_random_hand_gesture" />
    <wait time="500000" />
    <set_port name = "gesture" value = "%out_random_head_gesture" />
    <prompt>%input_dialog_string </prompt>
    <wait time="500000" />

    <set_port name = "input_dialog_string" value = "You could even seek help from your colleagues."/>
    <set_port name = "gesture" value = "%out_random_hand_gesture" />
    <set_port name = "return_gesture" value = "%out_random_hand_gesture" />
    <wait time="500000" />
    <set_port name = "gesture" value = "%out_random_head_gesture" />
    <prompt>%input_dialog_string </prompt>
        <wait time="500000" />

        <set_port name="robot_facial_expression" value="1" />
    <set_port name = "input_dialog_string" value = "Are you telling me you tried your best."/>
    <set_port name = "gesture" value = "%out_random_hand_gesture" />
    <set_port name = "return_gesture" value = "%out_random_hand_gesture" />
    <wait time="500000" />
    <set_port name = "gesture" value = "%out_random_head_gesture" />
    <prompt>%input_dialog_string </prompt>

    <set_port name="robot_facial_expression" value="1" />
    <set_port name = "input_dialog_string" value = "But still you could not achieve even 10 percent of the task assigned?"/>
    <set_port name = "gesture" value = "%out_random_hand_gesture" />
    <set_port name = "return_gesture" value = "%out_random_hand_gesture" />
    <wait time="500000" />
    <set_port name = "gesture" value = "%out_random_head_gesture" />
    <prompt>%input_dialog_string </prompt>

    <set_port name="robot_facial_expression" value="0" />
    <set_port name="activate_posture" value="1" />
    <set_port name="activate_facial_expressions" value="1" />
    <set_port name="head_pose_estimation" value="1" />
    <set_port name="get_shore" value="0" />
    <wait time="1000000" />

<mark name="check_3_answer" />
    <if port="out_finished_speaking" value="1" relation="==">""<goto mark="part_3" /></if>
    <elseif port="robot_static" value="1" relation="==">""<goto mark="perform_gesture_2" /></elseif>
    <else>""<goto mark="check_3_answer" /></else>

<mark name="perform_gesture_2" />
    <set_port name = "gesture" value = "random_open_down_up" />
    <wait time="500000" />
    <set_port name = "gesture" value = "random_head_nodding" />
    <goto mark="check_3_answer" />

<mark name="part_3" />
    <set_port name="activate_posture" value="0" />
    <set_port name="activate_facial_expressions" value="0" />
    <set_port name="head_pose_estimation" value="0" />
    <set_port name="get_shore" value="0" />

    <set_port name="robot_facial_expression" value="0" />
    <set_port name = "input_dialog_string" value = "You know, you used to be one of our sharpest employees."/>
    <set_port name = "gesture" value = "%out_random_hand_gesture" />
    <set_port name = "return_gesture" value = "%out_random_hand_gesture" />
    <wait time="500000" />
    <set_port name = "gesture" value = "%out_random_head_gesture" />
    <prompt>%input_dialog_string </prompt>
    <wait time="500000" />

    <set_port name="robot_facial_expression" value="7" />
    <set_port name = "input_dialog_string" value = "I was even planning on giving you a better position if you had performed well."/>
    <set_port name = "gesture" value = "%out_random_hand_gesture" />
    <set_port name = "return_gesture" value = "%out_random_hand_gesture" />
    <wait time="500000" />
    <set_port name = "gesture" value = "%out_random_head_gesture" />
    <prompt>%input_dialog_string </prompt>
    <wait time="500000" />

    <set_port name="robot_facial_expression" value="0" />
    <set_port name = "input_dialog_string" value = "I also mentioned this to you when I assigned you this task as a motivation."/>
    <set_port name = "gesture" value = "%out_random_hand_gesture" />
    <set_port name = "return_gesture" value = "%out_random_hand_gesture" />
    <wait time="500000" />
    <set_port name = "gesture" value = "%out_random_head_gesture" />
    <prompt>%input_dialog_string </prompt>
    <wait time="500000" />

    <set_port name="robot_facial_expression" value="11" />
    <set_port name = "input_dialog_string" value = "So, you tell me, what exactly has happened lately?"/>
    <set_port name = "gesture" value = "%out_random_hand_gesture" />
    <set_port name = "return_gesture" value = "%out_random_hand_gesture" />
    <wait time="500000" />
    <set_port name = "gesture" value = "%out_random_head_gesture" />
    <prompt>%input_dialog_string </prompt>
```

```
    <wait time="500000" />

    <set_port name="robot_facial_expression" value="0" />
    <set_port name="activate_posture" value="1" />
    <set_port name="activate_facial_expressions" value="1" />
    <set_port name="head_pose_estimation" value="1" />
    <set_port name="get_shore" value="0" />
    <wait time="1000000" />

<mark name="check_4_answer" />
    <if port="out_finished_speaking" value="1" relation="==">""<goto mark="part_4" /></if>
    <elseif port="robot_static" value="1" relation="==">""<goto mark="perform_gesture_3" /></elseif>
    <else>""<goto mark="check_4_answer" /></else>

<mark name="perform_gesture_3" />
    <set_port name = "gesture" value = "random_hand_movements_3" />
    <wait time="500000" />
    <set_port name = "gesture" value = "random_look_up" />
    <goto mark="check_4_answer" />

<mark name="part_4" />
    <set_port name="activate_posture" value="0" />
    <set_port name="activate_facial_expressions" value="0" />
    <set_port name="head_pose_estimation" value="0" />
    <set_port name="get_shore" value="0" />

    <set_port name="robot_facial_expression" value="17" />
    <set_port name = "input_dialog_string" value = "Listen, you are a good person, and I liked your
         work in the beginning."/>
    <set_port name = "gesture" value = "%out_random_hand_gesture" />
    <set_port name = "return_gesture" value = "%out_random_hand_gesture" />
    <wait time="500000" />
    <set_port name = "gesture" value = "%out_random_head_gesture" />
    <prompt>%input_dialog_string </prompt>
    <wait time="500000" />

    <set_port name="robot_facial_expression" value="0" />
    <set_port name = "input_dialog_string" value = "You made time for people, you used to actually know
          how to listen, and you got things done."/>
    <set_port name = "gesture" value = "%out_random_hand_gesture" />
    <set_port name = "return_gesture" value = "%out_random_hand_gesture" />
    <wait time="500000" />
    <set_port name = "gesture" value = "%out_random_head_gesture" />
    <prompt>%input_dialog_string </prompt>
        wait time="500000" />

        set_port name="robot_facial_expression" value="1" />
    <set_port name = "input_dialog_string" value = "Now all I hear is how much time you spend shopping
         online and chatting on facebook."/>
    <set_port name = "gesture" value = "%out_random_hand_gesture" />
    <set_port name = "return_gesture" value = "%out_random_hand_gesture" />
    <wait time="500000" />
    <set_port name = "gesture" value = "%out_random_head_gesture" />
    <prompt>%input_dialog_string </prompt>
        wait time="500000" />

        set_port name="robot_facial_expression" value="10" />
    <set_port name = "input_dialog_string" value = "Is it true? Do you have any explanation?"/>
    <set_port name = "gesture" value = "%out_random_hand_gesture" />
    <set_port name = "return_gesture" value = "%out_random_hand_gesture" />
    <wait time="500000" />
    <set_port name = "gesture" value = "%out_random_head_gesture" />
    <prompt>%input_dialog_string </prompt>

    <set_port name="robot_facial_expression" value="0" />
    <set_port name="activate_posture" value="1" />
    <set_port name="activate_facial_expressions" value="1" />
    <set_port name="head_pose_estimation" value="1" />
    <set_port name="get_shore" value="0" />
    <wait time="1000000" />

<mark name="check_5_answer" />
    <if port="out_finished_speaking" value="1" relation="==">""<goto mark="part_5" /></if>
    <elseif port="robot_static" value="1" relation="==">""<goto mark="perform_gesture_4" /></elseif>
    <else>""<goto mark="check_5_answer" /></else>

<mark name="perform_gesture_4" />
    <set_port name = "gesture" value = "random_open_up_down" />
    <wait time="500000" />
    <set_port name = "gesture" value = "random_head_shaking" />
    <goto mark="check_5_answer" />

<mark name="part_5" />
    <set_port name="activate_posture" value="0" />
    <set_port name="activate_facial_expressions" value="0" />
    <set_port name="head_pose_estimation" value="0" />
    <set_port name="get_shore" value="0" />

    <set_port name="robot_facial_expression" value="7" />
    <set_port name = "input_dialog_string" value = "Okay that is enough. I am done here."/>
    <set_port name = "gesture" value = "%out_random_hand_gesture" />
    <set_port name = "return_gesture" value = "%out_random_hand_gesture" />
    <wait time="500000" />
    <set_port name = "gesture" value = "%out_random_head_gesture" />
    <prompt>%input_dialog_string </prompt>
    <wait time="500000" />
```

```
    <set_port name="robot_facial_expression" value="1" />
    <set_port name = "input_dialog_string" value = "Go and finish the task until afternoon."/>
    <set_port name = "gesture" value = "pointing_front" />
    <wait time="500000" />
    <set_port name = "gesture" value = "%out_random_head_gesture" />
    <prompt>%input_dialog_string </prompt>
    <wait time="500000" />

    <set_port name="robot_facial_expression" value="9" />
    <set_port name = "input_dialog_string" value = "And let me be very clear, next time I do not want to hear any excuses."/>
    <set_port name = "gesture" value = "scolding_right" />
    <prompt>%input_dialog_string </prompt>
    <set_port name="robot_facial_expression" value="0" />
    <wait time="1000000" />
    <goto mark="end" />

<mark name="upset" />
    <set_port name="gesture" value="head_down_upset" />
    <set_port name="robot_facial_expression" value="5" />
    <prompt>I think you are not interested to talk to me.</prompt>

    <mark name="end" />
    <prompt>Good Bye.</prompt>
    <set_port name="gesture" value="reset" />
    <set_port name="activate_posture" value="0" />
    <set_port name="head_pose_estimation" value="0" />
    <set_port name="activate_facial_expressions" value="0" />

                </SUI_dialog>
            </dialog>
    </platform>
</data>
```

D.2 Weather Scenario Dialogue

Listing D.2: Dialog description used in an interaction experiment described in Section 7.3.

```
<?xml version="1.0" encoding="UTF-8"?>
<data xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xsi:noNamespaceSchemaLocation="sources/cpp/libraries/dialog_system/etc/DoAmi_dialog_description_scheme.xsd">

    <platform type="desktop">


            <dialog dialog_id="weather_scenario">
                <SUI_dialog>

<mark name="start" />
    <set_port name="robot_facial_expression" value="0" />
    <if port="humans_exist" value="1" relation=">=">""<goto mark="initiate_interaction" /></if>
    <else>""<goto mark="start" /></else>

<mark name="initiate_interaction" />
    <set_port name = "gesture" value = "reactive_right_hand_greeting_static" />
    <goto mark="greeting" />

<mark name="check_the_distance_1" />
    <wait time="500000" />
    <set_port name = "gesture" value = "neutral" />
    <if port="interaction_partner_x" value="1600" relation=">=">""<goto mark="come_here" /></if>
    <else>""<goto mark="intro" /></else>

<mark name="check_the_distance_2" />
    <set_port name="robot_facial_expression" value="11" />
    <wait time="3000000" />
    <set_port name="robot_facial_expression" value="0" />
    <if port="interaction_partner_x" value="1600" relation=">=">""<goto mark="wait_for_distance" /></if>
    <else>""<goto mark="intro" /></else>

<mark name="wait_for_distance" />
        <set_port name="robot_facial_expression" value="13" />
    <wait time="5000000" />
    <if port="interaction_partner_x" value="1600" relation=">=">""<goto mark="upset" /></if>
    <else>""<goto mark="intro" /></else>

<mark name="intro" />
    <set_port name="emphasis_tragedy" value="1" />
    <set_port name = "input_dialog_string" value = "Have you noticed the change in the weather lately?" />
    <set_port name = "gesture" value = "%out_random_hand_gesture" />
    <set_port name = "return_gesture" value = "%out_random_hand_gesture" />
    <wait time="500000" />
    <set_port name = "gesture" value = "%out_random_head_gesture" />
    <prompt>%input_dialog_string </prompt>
    <wait time="500000" />
```

```
    <set_port name="emphasis_tragedy_strong" value="1" />
    <set_port name = "input_dialog_string" value = "It is so bizarre. I think afterall global warming is real."/>
    <set_port name = "gesture" value = "%out_random_hand_gesture" />
    <set_port name = "return_gesture" value = "%out_random_hand_gesture" />
    <wait time="500000" />
    <set_port name = "gesture" value = "%out_random_head_gesture" />
    <prompt>%input_dialog_string </prompt>
    <wait time="500000" />

    <set_port name="emphasis_strong_skeptic" value="1" />
    <set_port name = "input_dialog_string" value = "Do you like the weather? Please answer either by head nodding or shaking."/>
    <set_port name = "gesture" value = "%out_random_hand_gesture" />
    <set_port name = "return_gesture" value = "%out_random_hand_gesture" />
    <wait time="500000" />
    <set_port name = "gesture" value = "%out_random_head_gesture" />
    <prompt>%input_dialog_string </prompt>
    <wait time="500000" />

    <set_port name="head_pose_estimation" value="1" />
    <set_port name="robot_facial_expression" value="0" />

<mark name="check_1_answer" />
    <wait time="500000" />
    <if port="human_head_gesture" value="8" relation="==">"Why do you think this is a good weather? Please tell me in detail"<goto mark="nodding_part" /></if>
    <elseif port="human_head_gesture" value="9" relation="==">"Why do you think this is a good weather? Please tell me in detail"<goto mark="nodding_part" /></elseif>
    <!-- <elseif port="human_right_static_gesture" value="11" relation="==">"Perfect."<goto mark="part_1" /></elseif>
    <elseif port="human_right_static_gesture" value="13" relation="==">"Great."<goto mark="part_1" /></elseif>-->
    <elseif port="human_head_gesture" value="10" relation="==">"Why do you think this is not a good weather? Please tell me in detail"<goto mark="shaking_part" /></elseif>
    <elseif port="human_head_gesture" value="11" relation="==">"Why do you think this is not a good weather? Please tell me in detail"<goto mark="shaking_part" /></elseif>
    <!--<elseif port="human_right_static_gesture" value="12" relation="==">"Okay"<goto mark="part_1" /></elseif>-->
    <else>""<goto mark="check_1_answer" /></else>

<mark name="nodding_part" />
    <set_port name="robot_facial_expression" value="4" />
    <set_port name="activate_posture" value="1" />
    <set_port name="activate_facial_expressions" value="1" />
    <set_port name="head_pose_estimation" value="1" />
    <set_port name="get_shore" value="1" />
    <set_port name="activate_traits_detection" value="1" />
    <wait time="1000000" />

<mark name="check_2_answer" />
    <if port="out_finished_speaking" value="1" relation="==">""<goto mark="part_2" /></if>
    <elseif port="robot_static" value="1" relation="==">""<goto mark="perform_gesture_1" /></elseif>
    <else>""<goto mark="check_2_answer" /></else>

<mark name="perform_gesture_1" />
    <set_port name = "gesture" value = "random_open_down_up" />
    <wait time="500000" />
    <set_port name = "gesture" value = "random_head_nodding" />
    <goto mark="check_2_answer" />

<mark name="part_2" />
    <set_port name="robot_facial_expression" value="0" />
    <set_port name="activate_posture" value="0" />
    <set_port name="activate_facial_expressions" value="0" />
    <set_port name="head_pose_estimation" value="0" />
    <set_port name="get_shore" value="0" />
    <set_port name="activate_traits_detection" value="0" />
    <wait time="1000000" />

 <mark name="part_2_1" />
    <if port="extroversion_mehrabian" value="11" relation="==">""<goto mark="excited" /></if>
    <elseif port="achievement" value="9" relation="==">""<goto mark="excited" /></elseif>
    <elseif port="anxiety" value="27" relation="==">""<goto mark="bored" /></elseif>
    <elseif port="shyness" value="29" relation="==">""<goto mark="bored" /></elseif>
    <elseif port="shyness" value="0" relation="==">""<goto mark="bored" /></elseif>
    <elseif port="shyness" value="30" relation="==">""<goto mark="bored" /></elseif>
    <else>""<goto mark="part_2_1" /></else>

 <mark name="excited" />
    <set_port name="teasing_smile" value="1" />
    <set_port name = "input_dialog_string" value = "Ahh great. I can see that you are excited about the weather."/>
    <set_port name = "gesture" value = "%out_random_hand_gesture" />
    <set_port name = "return_gesture" value = "%out_random_hand_gesture" />
    <wait time="500000" />
    <set_port name = "gesture" value = "%out_random_head_gesture" />
    <prompt>%input_dialog_string </prompt>
    <wait time="500000" />

    <set_port name="thanks_spiritual" value="1" />
    <set_port name = "input_dialog_string" value = "At least someone in the group is happy."/>
    <set_port name = "gesture" value = "%out_random_hand_gesture" />
    <set_port name = "return_gesture" value = "%out_random_hand_gesture" />
    <wait time="500000" />
    <set_port name = "gesture" value = "%out_random_head_gesture" />
```

```
        <prompt>%input_dialog_string </prompt>
        <wait time="500000" />

        <goto mark="finish" />

    <mark name="bored" />
        <set_port name="robot_facial_expression" value="7" />
        <set_port name = "input_dialog_string" value = "You said that you like this weather. However you
            don't seem so enthusiastic about it."/>
        <set_port name = "gesture" value = "%out_random_hand_gesture" />
        <set_port name = "return_gesture" value = "%out_random_hand_gesture" />
        <wait time="500000" />
        <set_port name = "gesture" value = "%out_random_head_gesture" />
        <prompt>%input_dialog_string </prompt>
        <wait time="500000" />

        <set_port name="emphasis_empathic_happy" value="1" />
        <set_port name = "input_dialog_string" value = "Maybe something else is bothering you."/>
        <set_port name = "gesture" value = "%out_random_hand_gesture" />
        <set_port name = "return_gesture" value = "%out_random_hand_gesture" />
        <wait time="500000" />
        <set_port name = "gesture" value = "%out_random_head_gesture" />
        <prompt>%input_dialog_string </prompt>
        <wait time="500000" />

        <set_port name="fake_smile" value="1" />
        <set_port name = "input_dialog_string" value = "Whatever it is, I hope everything is good."/>
        <set_port name = "gesture" value = "%out_random_hand_gesture" />
        <set_port name = "return_gesture" value = "%out_random_hand_gesture" />
        <wait time="500000" />
        <set_port name = "gesture" value = "%out_random_head_gesture" />
        <prompt>%input_dialog_string </prompt>
        <wait time="500000" />

            <goto mark="finish" />

    <mark name="shaking_part" />
        <set_port name="skeptic_moderate" value="1" />
        <set_port name="activate_posture" value="1" />
        <set_port name="activate_facial_expressions" value="1" />
        <set_port name="head_pose_estimation" value="1" />
        <set_port name="get_shore" value="1" />
        <set_port name="activate_traits_detection" value="1" />
        <wait time="1000000" />

    <mark name="check_3_answer" />
        <if port="out_finished_speaking" value="1" relation="==">""<goto mark="part_3" /></if>
        <elseif port="robot_static" value="1" relation="==">""<goto mark="perform_gesture_2" /></elseif>
        <else>""<goto mark="check_3_answer" /></else>

    <mark name="perform_gesture_2" />
        <set_port name = "gesture" value = "random_open_down_up" />
        <wait time="500000" />
        <set_port name = "gesture" value = "random_head_nodding" />
        <goto mark="check_3_answer" />

    <mark name="part_3" />
        <set_port name="robot_facial_expression" value="0" />
        <set_port name="activate_posture" value="0" />
        <set_port name="activate_facial_expressions" value="0" />
        <set_port name="head_pose_estimation" value="0" />
        <set_port name="get_shore" value="0" />
        <set_port name="activate_traits_detection" value="0" />
        <wait time="1000000" />

    <mark name="part_3_1" />
        <if port="extroversion_mehrabian" value="11" relation="==">""<goto mark="excited_1" /></if>
        <elseif port="achievement" value="9" relation="==">""<goto mark="excited_1" /></elseif>
        <elseif port="anxiety" value="27" relation="==">""<goto mark="bored_1" /></elseif>
        <elseif port="shyness" value="29" relation="==">""<goto mark="bored_1" /></elseif>
        <elseif port="shyness" value="0" relation="==">""<goto mark="bored_1" /></elseif>
        <elseif port="shyness" value="30" relation="==">""<goto mark="bored_1" /></elseif>
        <else>""<goto mark="part_3_1" /></else>

    <mark name="excited_1" />
        <set_port name="teasing_smile" value="1" />
        <set_port name = "input_dialog_string" value = "Ahh okay. But it seems that you are not effected by
            the weather at all."/>
        <set_port name = "gesture" value = "%out_random_hand_gesture" />
        <set_port name = "return_gesture" value = "%out_random_hand_gesture" />
        <set_port name = "gesture" value = "%out_primitive_gesture" />
        <wait time="500000" />
        <set_port name = "gesture" value = "%out_random_head_gesture" />
        <prompt>%input_dialog_string</prompt>
        <wait time="500000" />

        <set_port name="robot_facial_expression" value="4" />
        <set_port name = "input_dialog_string" value = "It is good to see you in high spirits today."/>
        <set_port name = "gesture" value = "%out_random_hand_gesture" />
        <set_port name = "return_gesture" value = "%out_random_hand_gesture" />
        <wait time="500000" />
        <set_port name = "gesture" value = "%out_random_head_gesture" />
        <prompt>%input_dialog_string </prompt>
        <wait time="500000" />

        <goto mark="finish" />
```

```

 <mark name="bored_1" />
    <set_port name="contempt_moderate" value="1" />
    <set_port name = "input_dialog_string" value = "Well I can understand totally your depressing behaviour."/>
    <set_port name = "gesture" value = "%out_random_hand_gesture" />
    <set_port name = "return_gesture" value = "%out_random_hand_gesture" />
    <set_port name = "gesture" value = "%out_primitive_gesture" />
    <wait time="500000" />
    <set_port name = "gesture" value = "%out_random_head_gesture" />
    <prompt>%input_dialog_string</prompt>
    <wait time="500000" />

    <set_port name="depressed_giving_info" value="1" />
    <set_port name = "input_dialog_string" value = "This weather is really terrible."/>
    <set_port name = "gesture" value = "%out_random_hand_gesture" />
    <set_port name = "return_gesture" value = "%out_random_hand_gesture" />
    <wait time="500000" />
    <set_port name = "gesture" value = "%out_random_head_gesture" />
    <prompt>%input_dialog_string </prompt>
    <wait time="500000" />

    <set_port name="pleasant_male" value="1" />
    <set_port name = "input_dialog_string" value = "Anyway cheer up now because the weather is going to change next week."/>
    <set_port name = "gesture" value = "%out_random_hand_gesture" />
    <set_port name = "return_gesture" value = "%out_random_hand_gesture" />
    <wait time="500000" />
    <set_port name = "gesture" value = "%out_random_head_gesture" />
    <prompt>%input_dialog_string </prompt>
    <wait time="500000" />

    <goto mark="finish" />

 <mark name="finish" />
    <set_port name="anger_moderate" value="1" />
    <set_port name = "input_dialog_string" value = "Oh shit. My battery is running out. I have to go and rest."/>
    <set_port name = "gesture" value = "%out_random_hand_gesture" />
    <set_port name = "return_gesture" value = "%out_random_hand_gesture" />
    <wait time="500000" />
    <set_port name = "gesture" value = "%out_random_head_gesture" />
    <prompt>%input_dialog_string </prompt>
    <wait time="500000" />
    <goto mark="end" />

 <mark name="greeting" />
    <set_port name="robot_facial_expression" value="4" />
    <set_variable name="rnd" random="9" />
    <if var="rnd" value="0" relation="==">"Hello"<goto mark="check_the_distance_1" /></if>
    <elseif var="rnd" value="1" relation="==">"Hey, is that you?"<goto mark="check_the_distance_1" /></elseif>
    <elseif var="rnd" value="2" relation="==">"Hi, how are you?"<goto mark="check_the_distance_1" /></elseif>
    <elseif var="rnd" value="3" relation="==">"Hey"<goto mark="check_the_distance_1" /></elseif>
    <elseif var="rnd" value="4" relation="==">"What's up"<goto mark="check_the_distance_1" /></elseif>
    <elseif var="rnd" value="6" relation="==">"Good to see you"<goto mark="check_the_distance_1" /></elseif>
    <elseif var="rnd" value="7" relation="==">"Nice to see you"<goto mark="check_the_distance_1" /></elseif>
    <elseif var="rnd" value="8" relation="==">"How's your day going"<goto mark="check_the_distance_1" /></elseif>
    <elseif var="rnd" value="9" relation="==">"How are you doing"<goto mark="check_the_distance_1" /></elseif>
    <else>""<goto mark="greeting" /></else>

<mark name="come_here" />
    <set_port name="casual_neutral" value="1" />
    <set_variable name="rnd" random="3" />
    <if var="rnd" value="0" relation="==">"Can you come here please?"<goto mark="check_the_distance_2" /></if>
    <elseif var="rnd" value="1" relation="==">"Come closer please."<goto mark="check_the_distance_2" /></elseif>
    <elseif var="rnd" value="2" relation="==">"Don't be afraid, come closer."<goto mark="check_the_distance_2" /></elseif>
    <elseif var="rnd" value="3" relation="==">" Come here, don't be shy."<goto mark="check_the_distance_2" /></elseif>
    <else>""<goto mark="come_here" /></else>

<mark name="upset" />
    <set_port name="gesture" value="head_down_upset" />
    <set_port name="emphasis_strong_catastrophic" value="1" />
    <prompt>I think you are not interested to talk to me.</prompt>

<mark name="end" />
        <set_port name="robot_facial_expression" value="4" />
    <prompt>Good Bye.</prompt>
    <set_port name="gesture" value="reset" />
    <set_port name="activate_posture" value="0" />
    <set_port name="head_pose_estimation" value="0" />
    <set_port name="activate_facial_expressions" value="0" />

                </SUI_dialog>
            </dialog>
    </platform>
</data>
```

E. Questionnaire

Q. No. 1)

Did you understand the topic of communication between Human and a Robin?

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
☐	☐	☐	☐	☐

Answer:

____

____

____

Q. No. 2)

What thing (gesture/expression/speech) did you like during the interaction? And Why?

- ____________________

Why?

____

____

____

- ____________________

Why?

____

____

____

Q. No. 3)

What thing (gesture/expression/speech) you don't like during the interaction? And Why?

- ____________________

Why?

____

____

____

- ____________________

Why?

____

____

____

Q. No. 4)

Can you write some emotions that Robin expresses during the interaction?

- ____________________
- ____________________
- ____________________
- ____________________
- ____________________
- ____________________
- ____________________
- ____________________
- ____________________
- ____________________

Q. No. 5)

Please write some comments about the mimics? (For example, robin was smiling most of the times, robin make crazy expressions in the start, etc.)

Answer:

____

____

____

Q. No. 6)

Has Robin expressed mimics and gestures naturally according to the situation? Why?

Yes ☐ No ☐

Answer:

Q. No. 7)

What do you think about the Robin's responses? (fast, slow, etc.)

Answer:

Q. No. 8)

Does Robin interact with a human in a human-like way? (on the scale of 1-10)

1 ☐—☐—☐—☐—☐—☐—☐—☐—☐—☐ 10

Q. No. 9)

Are you able to understand the Robin's speech during the interaction? If no, then why?

Answer:

Q. No. 10)

Please write some remarks about the whole interaction. (For example, the interaction seems natural, the interaction can be made better by ...., etc.)

Answer:

____

____

____

Bibliography

[Agarwal 04] A. Agarwal, B. Triggs, "3D human pose from silhouettes by relevance vector regression", in *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.*, vol. 2. June 2004, pp. 882–888 Vol.2.

[Al-Darraji 16a] S. Al-Darraji, K. Berns, A. Rodić, "Action Unit Based Facial Expression Recognition Using Deep Learning", in *Proceedings of the 25th International Conference on Robotics in Alpe-Adria-Danube Region, RAAD 2016*, ser. Advances in Intelligent Systems and Computing, A. Rodić, T. Borangiu, Eds., vol. 540. Cham, Switzerland: Springer International Publishing, July 1–2 2016, pp. 413–420. `http://www.springer.com/gp/book/9783319490571` ISBN: 978-3-319-49057-1.

[Al-Darraji 16b] S. Al-Darraji, Z. Zafar, K. Berns, "Real-time Perception of Nonverbal Human Feedback in a Gaming Scenario", in *Proceedings of the 2016 British HCI Conference*. Bournemouth, UK, July 11–15 2016.

[Al-Shawaf 16] L. Al-Shawaf, D. Conroy-Beam, K. Asao, D. M. Buss, "Human Emotions: An Evolutionary Psychological Perspective", *Emotion Review*, vol. 8, no. 2, pp. 173–186, 2016.

[Aly 13] A. Aly, A. Tapus, "A model for synthesizing a combined verbal and nonverbal behavior based on personality traits in human-robot interaction", in *Human-Robot Interaction (HRI), 2013 8th ACM/IEEE International Conference on*. March 2013, pp. 325–332.

[Ambridge 14] B. Ambridge, *Psy-Q: You know your IQ-now test your psychological intelligence*, Profile Books, 2014.

[Anderson 91] A. H. Anderson, M. Bader, E. G. Bard, E. Boyle, G. Doherty, S. Garrod, S. Isard, J. Kowtko, J. McAllister, J. Miller et al, "The HCRC map task corpus", *Language and speech*, vol. 34, no. 4, pp. 351–366, 1991.

[Argyle 13] M. Argyle, *Bodily communication*, Routledge, 2013.

[Barrick 91] M. R. Barrick, M. K. Mount, "The big five personality dimensions and job performance: a meta-analysis", *Personnel psychology*, vol. 44, no. 1, pp. 1–26, 1991.

[Bartneck 07] C. Bartneck, M. V. D. Hoek, O. Mubin, A. A. Mahmud, "Daisy, Daisy, give me your answer do!: switching off a robot", in *Proceedings of the ACM/IEEE international conference on Human-robot interaction*, ACM. 2007, pp. 217–222.

[Batrinca 11] L. M. Batrinca, N. Mana, B. Lepri, F. Pianesi, N. Sebe, “Please, tell me about yourself: automatic personality assessment using short self-presentations”, in *Proceedings of the 13th international conference on multimodal interfaces*, ACM. 2011, pp. 255–262.

[Batrinca 12] L. Batrinca, B. Lepri, N. Mana, F. Pianesi, “Multimodal recognition of personality traits in human-computer collaborative tasks”, in *Proceedings of the 14th ACM international conference on multimodal interaction*, ACM. 2012, pp. 39–46.

[Beer 14] J. M. Beer, A. D. Fisk, W. A. Rogers, “Toward a framework for levels of robot autonomy in human-robot interaction”, *Journal of human-robot interaction*, vol. 3, no. 2, pp. 74–99, 2014.

[Berns 09] K. Berns, N. Schmitz, “Perception System for Naturally Interacting Humanoid Robots”, *Künstliche Intelligenz (KI)*, vol. 1/09, 2009.

[Berns 18] K. Berns, Z. Zafar, “Emotion Based Human-Robot Interaction”, in *Proceedings of the 13th International Scientific-Technical Conference on Electromechanics and Robotics "Zavalishin’s Readings" (ER(ZR))*. St. Petersburg, Russia, April 18–21 2018.

[Biel 10] J.-I. Biel, D. Gatica-Perez, “Voices of vlogging”, in *Fourth International AAAI Conference on Weblogs and Social Media*. 2010.

[Biel 12] J.-I. Biel, D. Gatica-Perez, “The youtube lens: Crowdsourced personality impressions and audiovisual analysis of vlogs”, *IEEE Transactions on Multimedia*, vol. 15, no. 1, pp. 41–55, 2012.

[Boedeker 16] M. Boedeker, “Understanding Affective Experiences: Towards a Practical Framework in the VALIT-Project”, *TAMK English Journal*, March 4 2016.

[Boukricha 09] H. Boukricha, I. Wachsmuth, A. Hofst"atter, K. Grammer, “Pleasure-arousal-dominance driven facial expression simulation”, in *2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*, IEEE. 2009, pp. 1–7.

[Boureau 10] Y. L. Boureau, F. Bach, Y. LeCun, J. Ponce, “Learning mid-level features for recognition”, in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. June 2010, pp. 2559–2566.

[Brščić 13] D. Brščić, T. Kanda, T. Ikeda, T. Miyashita, “Person tracking in large public spaces using 3-D range sensors”, *IEEE Transactions on Human-Machine Systems*, vol. 43, no. 6, pp. 522–534, 2013.

[Bruyn 03] B. D. Bruyn, “Review: The History of Psychology: Fundamental Questions”, *Perception*, vol. 32, no. 11, pp. 1409–1410, 2003.

[Burgoon 89] J. K. Burgoon, D. B. Buller, W. G. Woodall, *Nonverbal communication: The unspoken dialogue*, Harpercollins College Division, 1989.

[Burke 04] J. L. Burke, R. R. Murphy, M. D. Coovert, D. L. Riddle, "Moonlight in Miami: Field study of human-robot interaction in the context of an urban search and rescue disaster response training exercise", *Human–Computer Interaction*, vol. 19, no. 1–2, pp. 85–116, 2004.

[Charles 08] S. T. Charles, M. Gatz, K. Kato, N. L. Pedersen, "Physical health 25 years later: the predictive ability of neuroticism.", *Health Psychology*, vol. 27, no. 3, p. 369, 2008.

[Chartrand 99] T. L. Chartrand, J. A. Bargh, "The chameleon effect: the perception–behavior link and social interaction.", *Journal of personality and social psychology*, vol. 76, no. 6, p. 893, 1999.

[Cherry 17] K. Cherry, "Erik Erikson's Stages of Psychosocial Development", *Psychology. Psychosocial Theories. P"aivitetty*, vol. 14, p. 2017, 2017.

[Christensen 05] H. I. Christensen, E. Pacchierotti, "Embodied social interaction for robots", *AISB-05*, pp. 40–45, 2005.

[Cohen 03] I. Cohen, "Noise Spectrum Estimation in Adverse Environments: Improved Minima Controlled Recurive Averaging", in *IEEE Transactions on Speech and Audio Processing*, vol. XX, no. Y. 2003.

[Costa 76] P. T. Costa, R. R. McCrae, "Age differences in personality structure: A cluster analytic approach", *Journal of gerontology*, vol. 31, no. 5, pp. 564–570, 1976.

[Costa 85] P. T. Costa, R. R. McCrae, *The NEO personality inventory*, Psychological Assessment Resources Odessa, FL, 1985.

[Costa 92] P. T. Costa, R. R. McCrae, "Normal personality assessment in clinical practice: The NEO Personality Inventory.", *Psychological assessment*, vol. 4, no. 1, p.$\tilde{5}$, 1992.

[Costa 98] P. T. Costa, R. R. McCrae, "Trait theories of personality", in *Advanced personality*, Springer, 1998, pp. 103–121.

[Dael 12] N. Dael, M. Mortillaro, K. R. Scherer, "Emotion expression in body action and posture.", *Emotion*, vol. 12, no. 5, p. 1085, 2012.

[Dautenhahn 13] K. Dautenhahn, "Human-Robot Interaction", in *The Encyclopedia of Human-Computer Interaction*, 2nd ed., M. Soegaard, R. F. e. . Dam, Eds., 2013.

[de Vries 13] R. E. de Vries, A. Bakker-Pieper, F. E. Konings, B. Schouten, "The communication styles inventory (CSI) a six-dimensional behavioral model of communication styles and its relation with personality", *Communication Research*, vol. 40, no. 4, pp. 506–532, 2013.

[DeGroot 09] T. DeGroot, J. Gooty, "Can nonverbal cues be used to make meaningful personality attributions in employment interviews?", *Journal of Business and Psychology*, vol. 24, no. 2, pp. 179–192, 2009.

[Dolan 06] S. Dolan, *Stress, self-esteem, health and work*, Springer, 2006.

[Ekman 69] P. Ekman, W. V. Friesen, “Nonverbal leakage and clues to deception”, *Psychiatry*, vol. 32, no. 1, pp. 88–106, 1969.

[Escalante 17] H. J. Escalante, I. Guyon, S. Escalera, J. Jacques, M. Madadi, X. Baró, S. Ayache, E. Viegas, Y. G"uçl"ut"urk, U. Guclu et al, “Design of an explainable machine learning challenge for video interviews”, in *IJCNN 2017-30th International Joint Conference on Neural Networks.* 2017, pp. 1–8.

[Eyben 10] F. Eyben, M. W"ollmer, B. Schuller, “Opensmile: the munich versatile and fast open-source audio feature extractor”, in *Proceedings of the 18th ACM international conference on Multimedia*, ACM. 2010, pp. 1459–1462.

[Eysenck 52] H. J. Eysenck, *The scientific study of personality.*, Macmillan, 1952.

[Eysenck 64] H. J. Eysenck et al, *Manual of the Eysenck personality inventory*, University of London Press London, 1964.

[Eysenck 67] H. J. Eysenck, S. Rachman, *The causes and cures of neurosis: An introduction to modern behaviour therapy based on learning theory and the principles of conditioning*, Taylor and Francis, 1967.

[Eysenck 85] H. Eysenck, M. Eysenck, “Personality and individual differences: a natural science approachPlenum”, *New York*, 1985.

[Faust 09] F. Faust, “Multi-Channel Voice based Speaker Recognition for Humanoid Robots”, Diploma thesis, Robotics Research Lab, Department of Computer Sciences, University of Kaiserslautern, October 2009. unpublished.

[Feil-Seifer 05] D. Feil-Seifer, M. J. Matarić, “Defining Socially Assistive Robotics”, in *9th International Conference on Rehabilitation Robotics.* Chicago, IL, USA, June 2005, pp. 465–468.

[Ferwerda 16] B. Ferwerda, M. Schedl, M. Tkalcic, “Using instagram picture features to predict users’ personality”, in *International Conference on Multimedia Modeling*, Springer. 2016, pp. 850–861.

[Fiske 10] S. T. Fiske, D. T. Gilbert, G. Lindzey, *Handbook of social psychology*, John Wiley andSons, 2010, vol. 2.

[Fong 03] T. Fong, I. Nourbakhsh, K. Dautenhahn, “A survey of socially interactive robots”, *Robotics and Autonomous Systems*, vol. 42, pp. 143–166, 2003.

[Freud 01] S. Freud, A. A. Brill, *Three contributions to the theory of sex*, Courier Corporation, 2001.

[Freud 54] S. Freud, P. M. Bonaparte, *The origins of psychoanalysis*, Imago London, 1954, vol. 216.

[Friedman 16] H. S. Friedman, M. W. Schustack, *Personality: Classic theories and modern research*, Pearson, 2016.

[Gardecki 17] A. Gardecki, M. Podpora, “Experience from the operation of the Pepper humanoid robots”, in *2017 Progress in Applied Electrical Engineering (PAEE)*. June 2017, pp. 1–6.

[Gardecki 18] A. Gardecki, M. Podpora, R. Beniak, B. Klin, “The Pepper Humanoid Robot in Front Desk Application”, in *2018 Progress in Applied Electrical Engineering (PAEE)*. June 2018, pp. 1–7.

[Ge 16] L. Ge, H. Tang, Q. Zhou, Y. Tang, J. Lang, “Classification Algorithms to Predict Students’ Extraversion-Introversion Traits”, in *Cyberworlds (CW), 2016 International Conference on*, IEEE. 2016, pp. 135–138.

[Gershgorn 17] D. Gershgorn, “Inside the mechanical brain of the world’s first robot citizen”, *Quartz, November*, vol. 16, 2017.

[Glas 09] D. F. Glas, T. Miyashita, H. Ishiguro, N. Hagita, “Laser-based tracking of human position and orientation using parametric shape modeling”, *Advanced robotics*, vol. 23, no. 4, pp. 405–428, 2009.

[Glas 16] D. F. Glas, T. Minato, C. T. Ishi, T. Kawahara, H. Ishiguro, “Erica: The erato intelligent conversational android”, in *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, IEEE. 2016, pp. 22–29.

[Gorbova 17] J. Gorbova, I. L"usi, A. Litvin, G. Anbarjafari, “Automated Screening Of Job Candidate Based On Multimodal Video Processing”, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. 2017, pp. 29–35.

[Gosling 03] S. D. Gosling, P. J. Rentfrow, W. B. S. Jr, “A very brief measure of the Big-Five personality domains”, *Journal of Research in personality*, vol. 37, no. 6, pp. 504–528, 2003.

[Gouaillier 09] D. Gouaillier, V. Hugel, P. Blazevic, C. Kilner, J. Monceaux, P. Lafourcade, B. Marnier, J. Serre, B. Maisonnier, “Mechatronic design of NAO humanoid”, in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*. Kobe, Japan, May 12–17 2009, pp. 769–774.

[Gu 04] H. Gu, Q. Ji, “An automated face reader for fatigue detection”, in *Sixth IEEE International Conference on Automatic Face and Gesture Recognition, 2004. Proceedings.*, IEEE. 2004, pp. 111–116.

[Hall 63] E. T. Hall, “A System for the notation of proxemic behavior 1”, *American anthropologist*, vol. 65, no. 5, pp. 1003–1026, 1963.

[Hargie 16] O. Hargie, *Skilled interpersonal communication: Research, theory and practice*, Routledge, 2016.

[Henderson 82] M. Henderson, A. Furnham, “Similarity and attraction: The relationship between personality, beliefs, skills, needs and friendship choice”, *Journal of Adolescence*, vol. 5, no. 2, pp. 111–123, 1982.

[Hirth 11] J. Hirth, N. Schmitz, K. Berns, "Towards Social Robots: Designing an Emotion-based Architecture", *International Journal of Social Robotics*, vol. 3, no. 3, pp. 273–290, 2011. DOI: 10.1007/s12369-010-0087-2.

[Hirth 12] J. Hirth, N. Schmitz, K. Berns, "Playing Tangram with the Humanoid Robot ROMAN", in *Proceedings of the German Conference on Robotics (ROBOTIK)*. Munich, Germany, May 21–22 2012, pp. 18–23.

[Hjelmas 01] E. Hjelmas, B. K. Low, "Face detection: A survey", in *Computer Vision and Image Understanding*, vol. 83, no. 3. Sept. 2001, pp. 236–274.

[Holt 89] R. R. Holt, *Freud reappraised: A fresh look at psychoanalytic theory.*, Guilford Press, 1989.

[Huang 11] Y. Huang, D. Monekosso, H. Wang, J. C. Augusto, "A Concept Grounding Approach for Glove-Based Gesture Recognition", in *Intelligent Environments (IE), 2011 7th International Conference on*. July 2011, pp. 358–361.

[Inoue 16] K. Inoue, P. Milhorat, D. Lala, T. Zhao, T. Kawahara, "Talking with ERICA, an autonomous android", in *Proceedings of the 17th Annual Meeting of the Special Interest Group on Discourse and Dialogue*. 2016, pp. 212–215.

[Isbister 00] K. Isbister, C. Nass, "Consistency of personality in interactive characters: verbal cues, non-verbal cues, and user characteristics", *International journal of human-computer studies*, vol. 53, no. 2, pp. 251–267, 2000.

[Jensen 16] M. Jensen, "Personality traits and nonverbal communication patterns", *Int'l J. Soc. Sci. Stud.*, vol. 4, p.$\tilde{5}$7, 2016.

[Jeronimus 14] B. F. Jeronimus, H. Riese, R. Sanderman, J. Ormel, "Mutual reinforcement between neuroticism and life experiences: A five-wave, 16-year study to test reciprocal causation.", *Journal of personality and social psychology*, vol. 107, no. 4, p. 751, 2014.

[Ji 06] Q. Ji, P. Lan, C. Looney, "A probabilistic framework for modeling and real-time monitoring human fatigue", *IEEE Transactions on systems, man, and cybernetics-Part A: Systems and humans*, vol. 36, no. 5, pp. 862–875, 2006.

[Joachim 52] H. H. Joachim, D. A. Rees, "Aristotle: The Nicomachean Ethics", *Journal of Philosophy*, vol. 49, no. 14, pp. 484–485, 1952.

[Joseph 18] A. Joseph, B. Christian, A. A. Abiodun, F. Oyawale, "A review on humanoid robotics in healthcare", *MATEC Web Conf.*, vol. 153, p. 02004, 2018.

[Joshi 14] J. Joshi, H. Gunes, R. Goecke, "Automatic prediction of perceived traits using visual cues under varied situational context", in *Pattern Recognition (ICPR), 2014 22nd International Conference on*, IEEE. 2014, pp. 2855–2860.

[Judge 99] T. A. Judge, C. A. Higgins, C. J. Thoresen, M. R. Barrick, "The big five personality traits, general mental ability, and career success across the life span", *Personnel psychology*, vol. 52, no. 3, pp. 621–652, 1999.

[Jung 71] C. G. Jung, *Psychological types, volume 6 of The collected works of CG Jung*, Princeton University Press, 1971.

[Kagan 09] J. Kagan, N. Snidman, *The long shadow of temperament*, Harvard University Press, 2009.

[Kaliouby 05] R. E. Kaliouby, P. Robinson, "Real-time inference of complex mental states from facial expressions and head gestures", in *Real-time vision for human-computer interaction*, Springer, 2005, pp. 181–200.

[Kanade 00] T. Kanade, J. F. Cohn, Y. Tian, "Comprehensive database for facial expression analysis", in *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580)*, IEEE. 2000, pp. 46–53.

[Kapoor 05] A. Kapoor, R. W. Picard, "Multimodal affect recognition in learning environments", in *Proceedings of the 13th annual ACM international conference on Multimedia*, ACM. 2005, pp. 677–682.

[Kapoor 07] A. Kapoor, W. Burleson, R. W. Picard, "Automatic prediction of frustration", *International journal of human-computer studies*, vol. 65, no. 8, pp. 724–736, 2007.

[Keltner 97] D. Keltner, "Facial expression, personality, and psychopathology", *What the face reveals*, pp. 450–452, 1997.

[Kiesler 04] S. Kiesler, P. Hinds, "Introduction to this special issue on human-robot interaction", *Human–Computer Interaction*, vol. 19, no. 1–2, pp. 1–8, 2004.

[Knapp 11] M. L. Knapp, J. A. Daly, *The SAGE handbook of interpersonal communication*, Sage Publications, 2011.

[Knapp 13] M. L. Knapp, J. A. Hall, T. G. Horgan, *Nonverbal Communication in Human Interaction*, Cengage Learning, 2013.

[Kobayashi 11] Y. Kobayashi, Y. Kuno, M. Gyoda, K. Yamazaki, Y. Seki, T. Tabata, M. Shibuya, "Assisted-care robot dealing with multiple requests in multi-party settings", in *2011 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. March 2011, pp. 167–168.

[Koch 07] J. Koch, M. Anastasopoulos, K. Berns, "Using the Modular Controller Architecture (MCA) in Ambient Assisted Living", in *3rd IET International Conference on Intelligent Environments (IE07)*. Ulm, Germany, September 24–25 2007, pp. 411–415.

[Kuhnke 12] E. Kuhnke, *Body language for dummies*, John Wiley & Sons, 2012.

[La-France 04] B. H. La-France, A. D. Heisel, M. J. Beatty, "Is there empirical evidence for a nonverbal profile of extraversion?: a meta-analysis and critique of the literature", *Communication Monographs*, vol. 71, no. 1, pp. 28–48, 2004.

[Laney 02] M. O. Laney, *The introvert advantage: How to thrive in an extrovert world*, Workman Publishing, 2002.

[Lee 06a] K. M. Lee, W. Peng, S.-A. Jin, C. Yan, “Can robots manifest personality?: An empirical test of personality recognition, social responses, and social presence in human–robot interaction”, *Journal of communication*, vol. 56, no. 4, pp. 754–772, 2006.

[Lee 06b] M. W. Lee, I. Cohen, “A model-based approach for estimating human 3D poses in static images”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 6, pp. 905–916, June 2006.

[Lillo 17] I. Lillo, J. C. Niebles, A. Soto, “Sparse composition of body poses and atomic actions for human activity recognition in RGB-D videos”, *Image and Vision Computing*, vol. 59, pp. 63 – 75, 2017.

[Lim 04] B.-C. Lim, R. E. Ployhart, “Transformational leadership: relations to the five-factor model and team performance in typical and maximum contexts.”, *Journal of applied psychology*, vol. 89, no. 4, p. 610, 2004.

[Lindemann 99] M. Lindemann, *Medicine and society in early modern Europe*, Cambridge University Press, 1999.

[MacKinnon 62] D. W. MacKinnon, “The nature and nurture of creative talent.”, *American psychologist*, vol. 17, no. 7, p. 484, 1962.

[Malik 02] J. Malik, G. Mori, “Estimating Human Body Configurations Using Shape Context Matching”, in *Proceedings of the 7th European Conference on Computer Vision-Part III*, ser. ECCV ’02. Springer-Verlag, 2002, pp. 666–680.

[Mandler 52] G. Mandler, S. B. Sarason, “A study of anxiety and learning.”, *The Journal of Abnormal and Social Psychology*, vol. 47, no. 2, p. 166, 1952.

[Matsumoto 13] D. Matsumoto, M. G. Frank, H. S. Hwang, *Nonverbal communication: Science and applications: Science and applications*, Sage, 2013.

[Matthews 03] G. Matthews, I. J. Deary, M. C. Whiteman, *Personality traits*, Cambridge University Press, 2003.

[McCrae 96] R. R. McCrae, “Social consequences of experiential openness.”, *Psychological bulletin*, vol. 120, no. 3, p. 323, 1996.

[McCrae 99] R. R. McCrae, P. T. C. Jr, “A five-factor theory of personality”, *Handbook of personality: Theory and research*, vol. 2, pp. 139–153, 1999.

[Mehrabian 69] A. Mehrabian, “Significance of posture and position in the communication of attitude and status relationships”, *Psychological Bulliten*, vol. 71, no. 5, pp. 359–372, 1969.

[Mehrabian 71] A. Mehrabian, “Nonverbal betrayal of feeling.”, *Journal of Experimental Research in Personality*, 1971.

[Mehrabian 78] A. Mehrabian, “Measures of individual differences in temperament”, *Educational and Psychological Measurement*, vol. 38, no. 4, pp. 1105–1117, 1978.

[Mehrabian 86] A. Mehrabian, I. V. Nahum, V. Duke, “Individual difference correlates and measures of predisposition to obesity and to anorexia”, *Imagination, Cognition and Personality*, vol. 5, no. 4, pp. 339–355, 1986.

[Mehrabian 95] A. Mehrabian, “Framework for a comprehensive description and measurement of emotional states.”, *Genetic, social, and general psychology monographs*, 1995.

[Mehrabian 96] A. Mehrabian, “Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament”, *Current Psychology: Developmental, Learning, Personality, Social*, vol. 14, 1996.

[Montepare 99] J. Montepare, E. Koff, D. Zaitchik, M. Albert, “The Use of Body Movements and Gestures as Cues to Emotions in Younger and Older Adults”, *Journal of Nonverbal Behavior*, vol. 23, no. 2, pp. 133–152, Jun 1999.

[Myers 80] I. B. Myers, P. B. Myers, *Gifts Differing: Understanding Personality Type*, Davies-Black Publishing, 1980.

[Myers 88] I. B. Myers, M. H. McCaulley, *Myers-Briggs Type Indicator: MBTI.*, Consulting Psychologists Press, 1988.

[Nakadai 10] K. Nakadai, T. Takahashi, H. G. Okuno, H. Nakajima, Y. Hasegawa, H. Tsujino, “Design and Implementation of Robot Audition System ’HARK’ - Open Source Software for Listening to Three Simultaneous Speakers”, *Advanced Robotics*, vol. 24, no. 5–6, pp. 739–761, 2010.

[Nass 01] C. Nass, K. M. Lee, “Does computer-synthesized speech manifest personality? Experimental tests of recognition, similarity-attraction, and consistency-attraction.”, *Journal of experimental psychology: applied*, vol. 7, no. 3, p. 171, 2001.

[Niebles 10] J. C. Niebles, C.-W. Chen, L. Fei-Fei, “Modeling temporal structure of decomposable motion segments for activity classification”, in *European conference on computer vision*, Springer. 2010, pp. 392–405.

[Nirjon 14] S. Nirjon et al, “Kintense: A robust, accurate, real-time and evolving system for detecting aggressive actions from streaming 3D skeleton data”, in *2014 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. March 2014, pp. 2–10.

[Nishio 07] S. Nishio, H. Ishiguro, N. Hagita, “Geminoid: Teleoperated android of an existing person”, in *Humanoid robots: New developments*, IntechOpen, 2007.

[Noller 06] P. Noller, *The SAGE Handbook of Nonverbal Communication*, SAGE Publications, Inc, 2006, ch. Nonverbal Communication in Close Relationships., pp. 403–421.

[Norris 07] C. J. Norris, J. T. Larsen, J. T. Cacioppo, “Neuroticism is associated with larger and more prolonged electrodermal responses to emotionally evocative pictures”, *Psychophysiology*, vol. 44, no. 5, pp. 823–826, 2007.

[Oberzaucher 08] E. Oberzaucher, K. Grammer, “Everything is movement: on the nature of embodied communication”, *Embodied communication in humans and machines*, pp. 151–177, 2008.

[of Seville 57] I. of Seville, *De Natura Rerum*, Liber etymologiarum, 1857, ch. Traitéde la nature.

[Olakitan 11] O. O. Olakitan, “An examination of the impact of selected personality traits on the innovative behaviour of entrepreneurs in Nigeria”, *International Business and management*, vol. 3, no. 2, pp. 112–121, 2011.

[Ones 96] D. S. Ones, C. Viswesvaran, A. D. Reiss, “Role of social desirability in personality testing for personnel selection: The red herring.”, *Journal of Applied Psychology*, vol. 81, no. 6, p. 660, 1996.

[Oosterwijk 09] S. Oosterwijk, M. Rotteveel, A. H. Fischer, U. Hess, “Embodied emotion concepts: How generating words about pride and disappointment influences posture”, *European Journal of Social Psychology*, vol. 39, no. 3, pp. 457–466, 2009.

[Pedro 11a] M. O. C. Z. S. Pedro, R. S. d Baker, M. M. T. Rodrigo, “Detecting carelessness through contextual estimation of slip probabilities among students using an intelligent tutor for mathematics”, in *International Conference on Artificial Intelligence in Education*, Springer. 2011, pp. 304–311.

[Pedro 11b] M. O. C. Z. S. Pedro, M. M. T. Rodrigo, R. S. d Baker, “The relationship between carelessness and affect in a Cognitive Tutor”, in *International Conference on Affective Computing and Intelligent Interaction*, Springer. 2011, pp. 306–315.

[Ramanathan 19] M. Ramanathan, N. Mishra, N. M. Thalmann, “Nadine Humanoid Social Robotics Platform”, in *Computer Graphics International Conference*, Springer. 2019, pp. 490–496.

[Rappaport 63] S. R. Rappaport, “Pattern and growth in personality”, *Journal of the American Academy of Child Psychiatry*, vol. 2, no. 4, pp. 769–771, 1963.

[Reichardt 13] M. Reichardt, T. Föhst, K. Berns, “On Software Quality-motivated Design of a Real-time Framework for Complex Robot Control Systems”, in *Proceedings of the 7th International Workshop on Software Quality and Maintainability (SQM), in conjunction with the 17th European Conference on Software Maintenance and Reengineering (CSMR)*. Genoa, Italy, March 5 2013.

[Reichardt 17] M. Reichardt, S. Schütz, K. Berns, “One Fits More - On Highly Modular Quality-Driven Design of Robotic Frameworks and Middleware”, *Journal of Software Engineering for Robotics (JOSER): Special Issue on Robotic Computing*, vol. 8, no. 1, pp. 141–153, December 2017. ISSN: 2035-3928; this publication is available at `http://joser.unibg.it/index.php?journal=joser&page=issue&op=view&path[]=9`.

[Roccas 02] S. Roccas, L. Sagiv, S. H. Schwartz, A. Knafo, “The big five personality factors and personal values”, *Personality and social psychology bulletin*, vol. 28, no. 6, pp. 789–801, 2002.

[Rogers 46] C. R. Rogers, “Significant aspects of client-centered therapy”, *American Psychologist*, vol. 1, no. 10, pp. 415–422, 1946.

[Rossberg-Gempton 93] I. Rossberg-Gempton, G. D. Poole, “The effect of open and closed postures on pleasant and unpleasant emotions”, *The Arts in Psychotherapy*, vol. 20, no. 1, pp. 75 – 82, 1993. Special Issue Research in the Creative Arts Therapies.

[Rothmann 03] S. Rothmann, E. P. Coetzer, “The big five personality dimensions and job performance”, *SA Journal of Industrial Psychology*, vol. 29, no. 1, pp. 68–74, 2003.

[Ruch 94] W. Ruch, “Extraversion, alcohol, and enjoyment”, *Personality and Individual Differences*, vol. 16, no. 1, pp. 89–102, 1994.

[Russell 77] J. A. Russell, A. Mehrabian, “Evidence for a three-factor theory of emotions”, *Journal of research in Personality*, vol. 11, no. 3, pp. 273–294, 1977.

[Russell 80] J. A. Russell, “The circumplex model of affect”, *Journal of Personality and Social Psychology*, vol. 39, pp. 1161–1178, 1980.

[Salam 16] H. Salam, O. Celiktutan, I. Hupont, H. Gunes, M. Chetouani, “Fully automatic analysis of engagement and its relationship to personality in human-robot interactions”, *IEEE Access*, vol. 5, pp. 705–721, 2016.

[Salam 17] H. Salam, O. Celiktutan, I. Hupont, H. Gunes, M. Chetouani, “Fully automatic analysis of engagement and its relationship to personality in human-robot interactions”, *IEEE Access*, vol. 5, pp. 705–721, 2017.

[Saleh 13] S. Saleh, A. Kickton, J. Hirth, K. Berns, “Robust Perception of an Interaction Partner Using Depth Information”, in *Proceeding of the International Conference on Advances in Computer-Human Interactions (ACHI)*. Nice, France, February 24–March 1 2013.

[Saleh 15] S. Saleh, K. Berns, “Nonverbal Communication With a Humanoid Robot Via Head Gestures”, in *PETRA ’15: Proceedings of the 8th International Conference on PErvasive Technologies Related to Assistive Environments*. Corfu, Greece: ACM, New York, NY, USA, July 1–3 2015. ISBN 978-1-4503-3452-5/15/07.

[Sanchez-Cortes 12] D. Sanchez-Cortes, O. Aran, M. S. Mast, D. Gatica-Perez, “A nonverbal behavior approach to identify emergent leaders in small groups”, *IEEE Transactions on Multimedia*, vol. 14, no. 3, pp. 816–832, 2012.

[Sanfilippo 15] F. Sanfilippo, L. I. Hatledal, H. Zhang, M. Fago, K. Y. Pettersen, “Controlling Kuka Industrial Robots: Flexible Communication Interface JOpenShowVar”, *IEEE Robotics Automation Magazine*, vol. 22, no. 4, pp. 96–109, Dec 2015.

[Scherer 77] K. R. Scherer, U. Scherer, J. A. Hall, R. Rosenthal, “Differential attribution of personality based on multi-channel presentation of verbal and nonverbal cues”, *Psychological Research*, vol. 39, no. 3, pp. 221–247, 1977.

[Schmitz 09] N. Schmitz, C. Spranger, K. Berns, "3D Audio Perception System for Humanoid Robots", in *Proceedings of the International Conferences on Advances in Computer-Human Interactions (ACHI).* Cancun, Mexico, February 1–7 2009, pp. 181–186.

[Schmitz 11] N. Schmitz, K. Berns, "Perception Systems for Naturally Interacting Humanoid Robots", in *IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN).* Atlanta, GA, USA, July 31 – August 3 2011, pp. 425–432.

[Seo 08] M.-G. Seo, L. F. Barrett, S. Jin, *The structure of affect: History, theory, and implications for emotion research in organizations.*, Edward Elgar Publishing, 2008.

[Shigemi 18] S. Shigemi, "ASIMO and humanoid robot research at Honda", *Humanoid robotics: A reference*, pp. 1–36, 2018.

[Skinner 53] B. F. Skinner, *Science and human behavior*, Simon and Schuster, 1953.

[Soldz 99] S. Soldz, G. E. Vaillant, "The Big Five personality traits and the life course: A 45-year longitudinal study", *Journal of Research in Personality*, vol. 33, no. 2, pp. 208–232, 1999.

[Spielman 18] R. M. Spielman, W. Jenkins, K. Dumper, M. Lovett, M. Perlmutter et al, *Psychology*, OpenStax, Rice University, 2018, ch. Personality, pp. 365–404.

[Srivastava 12] R. Srivastava, J. Feng, S. Roy, S. Yan, T. Sim, "Don't ask me what i'm like, just watch and listen", in *Proceedings of the 20th ACM international conference on Multimedia*, ACM. 2012, pp. 329–338.

[Staiano 11] J. Staiano, B. Lepri, R. Subramanian, N. Sebe, F. Pianesi, "Automatic modeling of personality states in small group interactions", in *Proceedings of the 19th ACM international conference on Multimedia*, ACM. 2011, pp. 989–992.

[Tapus 07] A. Tapus, M. J. Matarić, B. Scassellati, "The Grand Challenges in Socially Assistive Robotics", *Robotics and Automation Magazin*, vol. 14, no. 1, pp. 35–42, 2007.

[Tapus 08] A. Tapus, M. J. Mataric, "Socially Assistive Robots: The Link between Personality, Empathy, Physiological Signals, and Task Performance.", in *AAAI spring symposium: emotion, personality, and social behavior.* 2008, pp. 133–140.

[Tarasenko 10] S. Tarasenko, "Emotionally colorful reflexive games", *arXiv preprint arXiv:1101.0820*, 2010.

[Thalmann 17] N. M. Thalmann, L. Tian, F. Yao, "Nadine: A social robot that can localize objects and grasp them in a human way", in *Frontiers in Electronic Technologies*, Springer, 2017, pp. 1–23.

[Toegel 12] G. Toegel, J.-L. Barsoux, "How to become a better leader", *MIT Sloan Management Review*, vol. 53, no. 3, p.$\tilde{5}$1, 2012.

[Tong 01] S. Tong, E. Chang, "Support Vector Machine Active Learning for Image Retrieval", in *Proceedings of the Ninth ACM International Conference on Multimedia*, ser. MULTIMEDIA '01. New York, NY, USA: ACM, 2001, pp. 107–118.

[Uleman 08] J. S. Uleman, S. A. Saribay, C. M. Gonzalez, “Spontaneous inferences, implicit impressions, and implicit theories”, *Annu. Rev. Psychol.*, vol. 59, pp. 329–360, 2008.

[Ventura 17] C. Ventura, D. Masip, A. Lapedriza, “Interpreting CNN Models for Apparent Personality Trait Regression”, in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2017 IEEE Conference on*, IEEE. 2017, pp. 1705–1713.

[Vinciarelli 14] A. Vinciarelli, G. Mohammadi, “A survey of personality computing”, *IEEE Transactions on Affective Computing*, vol. 5, no. 3, pp. 273–291, 2014.

[Viola 04] P. Viola, M. J. Jones, “Robust Real-Time Face Detection”, *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137–154, May 2004.

[Wang 16] J. Wang et al, “Wearable sensor based human posture recognition”, in *2016 IEEE International Conference on Big Data (Big Data)*. Dec 2016, pp. 3432–3438.

[Watson 85] D. Watson, A. Tellegen, “Toward a consensual structure of mood.”, *Psychological bulletin*, vol. 98, no. 2, p. 219, 1985.

[Watson 97] D. Watson, L. A. Clark, “Measurement and mismeasurement of mood: Recurrent and emergent issues”, *Journal of personality assessment*, vol. 68, no. 2, pp. 267–296, 1997.

[Weinberg 99] R. Weinberg, D. Gould, “Personality and sport”, *Foundations of sport and exercise psychology*, pp. 25–46, 1999.

[Weller 17] C. Weller, “Meet the first-ever robot citizen - a humanoid named Sophia that once said it would’destroy humans”’, *Business Insider*, pp. 7–10, 2017.

[White 61] A. R. White, “Carelessness, indifference and recklessness”, *The Modern Law Review*, vol. 24, no. 5, pp. 592–595, 1961.

[Woods 07] S. Woods, K. Dautenhahn, C. Kaouri, R. te Boekhorst, K. L. Koay, M. L. Walters, “Are robots like people?: Relationships between participant and robot personality traits in human–robot interaction studies”, *Interaction Studies*, vol. 8, no. 2, pp. 281–305, 2007.

[Woods 16] S. A. Woods, F. C. Patterson, A. Koczwara, J. A. Sofat, “The value of being a conscientious learner: Examining the effects of the Big Five personality traits on self-reported learning from training”, *Journal of Workplace Learning*, vol. 28, no. 7, pp. 424–434, 2016.

[Wundt 97] W. M. Wundt, C. H. Judd, *Outlines of psychology*, Scholarly Press, 1897, vol. 1.

[Yang 12] J.-Y. Yang, D.-S. Kwon, “The effect of multiple robot interaction on human-robot interaction”, in *2012 9th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*. Nov 2012, pp. 30–33.

[Yang 17] P. Yang, K. Sasaki, K. Suzuki, K. Kase, S. Sugano, T. Ogata, “Repeatable Folding Task by Humanoid Robot Worker Using Deep Learning”, *IEEE Robotics and Automation Letters*, vol. 2, no. 2, pp. 397–403, April 2017.

[Zafar 16a] Z. Zafar, K. Berns, "Recognizing Hand Gestures for Human-Robot Interaction", in *Proceedings of the 9th International Conference on Advances in Computer-Human Interactions (ACHI).* Venice, Italy, April 24–28 2016, pp. 333–338.

[Zafar 16b] Z. Zafar, K. Berns, A. Rodić, "Recognizing Hand Gestures Using Local Features: A Comparison Study", in *Advances in Intelligent Systems and Computing: Proceedings of the 25th International Conference on Robotics in Alpe-Adria-Danube Region, RAAD 2016*, ser. Mechanisms and Machine Science, A. Rodić, T. Borangiu, Eds., vol. 540. Cham, Switzerland: Springer International Publishing, June 30–July 2 2016, pp. 394–401. `http://www.springer.com/gp/book/9783319490571` ISBN: 978-3-319-49057-1.

[Zafar 17] Z. Zafar, D. A. S. Villarreal, S. Al-Darraji, D. Urukalo, B. Karsten, A. Rodić, "Human Robot Interaction using Dynamic Hand Gestures", in *Advances in Service and Industrial Robotics: Proceedings of the 26th International Conference on Robotics in Alpe-Adria-Danube Region, RAAD 2017*, ser. Mechanisms and Machine Science, C. Ferraresi, G. Quaglia, Eds., vol. 49. Cham, Switzerland: Springer International Publishing, June 21–23 2017, pp. 647–656. `http://www.springer.com/gp/book/9783319612751` ISBN: 978-3-319-61275-1.

[Zafar 18a] Z. Zafar, S. H. Paplu, K. Berns, "Automatic Assessment of Human Personality Traits: A Step towards Intelligent Human-Robot Interaction", in *2018 IEEE-RAS 18th International Conference on Humanoid Robotics (Humanoids).* Beijing, China, November 6–9 2018, pp. 670–675.

[Zafar 18b] Z. Zafar, S. H. Paplu, K. Berns, "Real-time Recognition of Extroversion-Introversion Trait in Context of Human-Robot Interaction", in *Advances in Service and Industrial Robotics: Proceedings of the 27th International Conference on Robotics in Alpe-Adria-Danube Region, RAAD 2018*, ser. Mechanisms and Machine Science, N. A. Aspragathos, P. Koustoumpardis, V. Moulianitis, Eds., vol. 67. Cham, Switzerland: Springer International Publishing, June 6–8 2018, pp. 63–70. `https://doi.org/10.1007/978-3-030-00232-9` ISBN: 978-3-030-00232-9.

[Zafar 18c] Z. Zafar, R. Venugopal, K. Berns, "Real-Time Recognition of Human Postures for Human-Robot Interaction", in *Proceedings of the 11th International Conference on Advances in Computer-Human Interactions (ACHI).* Rome, Italy, March 25–29 2018, pp. 114–119.

[Zen 10] G. Zen, B. Lepri, E. Ricci, O. Lanz, "Space speaks: towards socially and personality aware visual surveillance", in *Proceedings of the 1st ACM international workshop on Multimodal pervasive video analysis*, ACM. 2010, pp. 37–42.

[Zhang 14] Z. Zhang, Y. Liu, A. Li, M. Wang, "A novel method for user-defined human posture recognition using Kinect", in *2014 7th International Congress on Image and Signal Processing*, IEEE. 2014, pp. 736–740.

About this Book

This thesis contributes to the development of an emotionally intelligent social robot for human-robot interaction. The system assesses personality traits of an interaction partner during interaction to adapt the robot's behaviour. The personality trait system is organized in three separate levels: perceptual, affective and behaviour level.

Perceptual level is responsible for enabling the robot to perceive, recognize and understand human actions in the surrounding environment to make sense of the situation. Affective level helps the robot to connect the knowledge acquired in the first level to make higher order evaluations such as assessment of human personality traits. Behaviour level uses the information from perceptual and affective level to enable a social robot to synthesize an appropriate behaviour adapted to human personality.

About the Author

Zuhair Zafar studied Electrical Engineering from 2006 till 2011 at GC University Lahore, Pakistan. He got a Master's degree in Computer Engineering from Lahore University of Management Sciences in 2013. From 2014 till 2020, he conducted his PhD research regarding social robotics in the field of human-robot interaction at the Robotics Research Lab, the research group led by Prof. Dr. Karsten Berns. His research interests include robot perception systems in the context of human-robot interaction.