ACTION ESTIMATION USING A THEORY OF MIND AS APPLIED ON THE HUMANOID ROBOT SURALP

by
Selim Özel

Submitted to the Graduate School of Engineering and Natural Sciences in partial fulfillment of the requirements for the degree of Master of Science

Sabanci University
August 2013

ACTION ESTIMATION USING A THEORY OF MIND AS APPLIED TO HUMANOID ROBOT SURALP

APPROVED BY:

Assoc. Prof. Dr. Kemalettin ERBATUR ……………………………………..
(Thesis Advisor)

Assoc. Prof. Dr. Volkan PATOGLU ……………………………………..

Assoc. Prof. Dr. Albert LEVI ……………………………………..

Assoc. Prof. Dr. Serhat YEŞİLYURT ……………………………………..

Assist. Prof. Dr. Hakan ERDOĞAN ……………………………………..

DATE OF APPROVAL: ……………………………………..

ACTION ESTIMATION USING A THEORY OF MIND AS APPLIED TO HUMANOID ROBOT SURALP

Selim ÖZEL

Mechatronics Engineering, Ms. Thesis, 2013

Thesis Supervisor: Assoc. Prof. Dr. Kemalettin ERBATUR

Keywords: Human robot interaction, humanoid robot, action estimation, theory of mind, computational theory of mind, mirror neurons

ABSTRACT

Explanations regarding human consciousness have existed for a very long time. Theory of Mind (ToM) is one of the contemporary explanations for counciousness. This theory states that humans have functionalized brain parts for understanding beliefs and intentions of others. Humans have an inherent ability for making inferences on visual data once an acition is observed. Understanding/anticipating human actions based on visual data can be explained in context of ToM.

It is proposed that a functionalized brain part is used for estimating intentions of others from observed movements of an actor. This functionalized part posses a Forward Model (FM) which simulates consequences of intentions. Simulated intentions are compared with observed movements to estimate the action of the actor. This thesis is based on implementation of such an action estimation model on a humanoid robot platform.

A computational model for the part of the human brain which estimates intentions is needed to implement the model on a robotic platform. There is a proposed computational model in the literature for the part of the brain which estimates intentions. Model explains how a FM can be used along with a loop for action estimation by providing an algorithm.

Motivation for such an implementation has two main reasons: To program a humanoid robot platform in such a way that it anticipates movements of the human actor to assist him/her, and a platform which can test ToM related to action estimation. In thesis the

implementation is made on SURALP (**S**abanci **U**niversity **R**ese**A**rch **L**abaratory **P**latform). Kinect is used for visual data input device. Various tests, which observe capabilities and limitations of the computational model, are completed with success.

İnsansı Robot SURALP Üzerinde Uygulanan Zihin Teorisi Tabanlı Hareket Tahmini

Selim ÖZEL

Mekatronik Mühendisliği Programı, Master Tezi, 2013

Tez Danışmanı: Doç. Dr. Kemalettin ERBATUR

Anahtar Kelimeler: İnsan robot etkileşimi, insansı robot, hareket tahmini, zihin teorisi, zihin hesaplama teorisi, ayna nöronlar

ÖZET

İnsanın bilincine dair açıklamalar uzun zamandır yapılmaktadır. Zihin Teorisi bilincin ne olduğunu açıklamaya çalışan çağdaş teorilerden biri olup, insanların zihinlerinin başkalarının inanışlarını ve niyetlerini anlamakta uzmanlaşmış parçalardan oluştuğunu öne sürer. İnsanlar, hareketleri görsel olarak fark ettiklerinde bilinçli/bilinçsiz çıkarımlar yapmaya başlarlar. İnsan hareketlerinin görsel etkileri üzerinden haretin niyetini anlamaya çalışmak veya hareketin sonuçlarını önceden fark etmek, Zihin Teorisi kapsamında açıklanabilecek davranışlardır.

Gözlemlenen insan hareketlerinin bilgisi üzerinden, hareketi yapmakta olan kişinin niyetine ilişkin çıkarımların yapıldığı fonksiyonlaşmış bir beyin kısmının varolduğu önerilmektedir. Beynin bu kısmında bulunan bir İleri Model sayesinde insan zihnindeki niyetlerin sonuçlarının simülasyonu yapıldığı da ileri sürülmektedir. Simülasyonlardan elde edilen bilgi ile gözlem sonucu elde edilen bilginin karşılaştırılması sonucunda insan zihni karşıdaki aktörün niyetini tahmin edebilir. Bu tezin amacı önerilen zihinsel niyet tahmini modelinin, bir robot platformu üzerinde denenmesidir.

Niyet tahmini fikrinin robot platformuna aktarılabilmesi için hesaplanabilir bir model gerekmekte olup böyle bir model literatürde mevcuttur. Bu model İleri Model'in nasıl hesaplanabileceğini ve niyet tamini yapan bir algoritma ile birlikte nasıl çalışacağını açıklamaktadır.

Modelin robot platformu üzerinde uygulanması fikrinin iki ana sebebi var: İnsansı bir robotu insanların davranışlarını önceden tahmin edebileceği bir şekilde programlayarak insanlara yardımcı olmasını sağlamak ve Zihin Teorisi dahilinde önerilmiş fikirlerin test edilebileceği bir platform oluşturmak. Tezdeki uygulama için kullanılan robot SURALP (**S**abancı **Ü**niversitesi **R**obot **A**raştırmaları **L**aboratuvar **P**latformu). Hareket gözlemi esnasında veri toplamak için kullanılan kamera da Kinect'dir. Önerilen niyet tahmin modelinin becerilerini ve sınırlarını analiz etmek için yapılmış olan testlerden başarılı sonuçlar alınmıştır.

ACKNOWLEDGEMENTS

I would like to express my gratitude for my thesis advisor Assoc. Prof. Kemalettin Erbatur. Throughout my Master of Science education, he always encouraged and supported me. He guided me in my graduate education and on this thesis with great enthusiasm.

I would also like to state my appreciation and regards to my thesis jury members Assoc. Prof. Dr. Volkan Patoğlu, Assoc. Prof. Dr. Albert Levi, Assoc. Assoc. Prof. Dr. Serhat Yeşilyurt and Assist. Prof. Dr. Hakan Erdoğan, for pointing their valuable ideas.

My student colleagues Emre Eskimez, Kaan Can Fidan, Tunç Akbaş, Utku Seven, and Ömer Kemal Adak deserve particular thanks for their invaluable support and friendship.

I would like to thank Iyad Hashlamon, Beste Bahçeci, Alper Yıldırım, Sanem Evren, Taygun Kekeç, Soner Ulun, Mehmet Ali Güney Alper Yıldırım, Mine Saraç, Edin Golubovic, and many more friends from Mechatronics Engineering Program.

Finally and most importantly, I want to express my gratefulness to my family, Fatih, Ayşe and Sinan Özel for their invaluable love, caring and support throughout my life.

ACTION ESTIMATION USING A THEORY OF MIND AS APPLIED TO HUMANOID ROBOT SURALP

TABLE OF CONTENTS

ABSTRACT iv

ÖZET vi

ACKNOWLEDGEMENTS viii

TABLE OF CONTENTS.. ix

LIST OF FIGURES xi

LIST OF TABLES xiv

1. INTRODUCTION 1

2. A SURVEY ON HUMAN ROBOT INTERACTION 3

3. BASIS OF THESIS 8

3.1. Neurological Basis of Intentions Estimation 8

3.2. Technical Realization of Intention Estimation 11

4. HARDWARE USED IN THESIS 13

4.1. Kinect 13

4.2. SURALP: A Full Body Humanoid Robot 16

4.2.1. Kinematic Arrangement of Arms 17

4.2.2. Inverse Kinematics Based on Visual Data Obtained from Kinect.......................20

4.2.3. Actuation of Arms...23

5. A HUMANOID ROBOT PLATFORM CAPABLE OF UNDERSTANDING INTENTIONS OF OTHERS 24

5.1. Neurological Details of FM...26

5.2. Computational Details of FM...27

5.3. Modifications of Thesis to Proposed FM...33

5.4. Setting 35

5.5. Experiments & Results...40

6. CONCLUSIONS 47

6.1. Limitations of the FM in the Thesis..47

6.2. Possible Improvement on the FM...48

6.3. Future Work..49

REFERENCES...51

LIST OF FIGURES

Figure 2.1: Kismet, a robot which is capable of interacting with humans through understanding and mimicking emotions..4

Figure 2.2: SONY AIBO ERS7A 4

Figure 2.3: Philos, a social robot developed at Case Western Reserve University 5

Figure 2.4: Atlas, designed and produced by Boston Dynamics 6

Figure 2.5: BLEEX was developed by Berkeley Robotics and Human Engineering Laboratory...7

Figure 3.2.1: Cerebral cortex of human brain is shown. It is the outermost layer of neurons in the brain. Premotor cortex is located in frontal lobe. Cerebellum can be seen in the bottom. 12

Figure 4.1.1: In the coordinate system z represents the direction which camera is looking at, x is the direction to the left of the device, and y is perpendicular to x and z. 13

Figure 4.1.2: Two different skeletons captured by kinect are shown. Left is the skeleton of a standing person, right is the seated version. 14

Figure 4.1.3: Coordinate axis assignments of Kinect Shoulder frame is given along with neck and elbow...15

Figure 4.2.1: SURALP 16

Figure 4.2.2: The Dimensions of SURALP in milimeters 17

Figure 4.2.1.1: Kinematic Arrengement...18

Figure 4.2.1.2: Denavit-Hartenberg based axis assignment for 6-DoF arm. Green arrows denote y axes, red x axes and blue z axes. 19

Figure 4.2.3.1: Hardware Architecture of SURALP. 23

Figure 5.1.1: Block representation of the proposed FM for action estimation in human brain. Image is obtained from [008]. 26

Figure 5.2.1: Three different trajectories for the intention of moving hand above head level is shown. Each joint trajectory is represented by a color. Color representations are given in Table 5.2.1. 29

Figure 5.2.2: These plots show time differentitaion of joint angle trajectories of three different moves from Figure 5.2.1. Color representations are in Table 5.2.1. Notice that 5^{th}joint angle trajectories are omitted in time differentiations....................................32

Figure 5.3.1: Block Diagram Representation of the Proposed Intention Estimation Loop. // Denotes that SURALP's Control Computer is not utilized until a decision is made..35

Figure 5.4.1: Joint angle trajectories for intention of reaching with an elbow down configuration, intention (1) 36

Figure 5.4.2: Joint angle trajectories for intention of reaching with an elbow up configuration, intention (2) 37

Figure 5.4.3: Joint angle trajectories for intention of moving hand above head level, intention (3)..37

Figure 5.4.4: Elbow down intention in a real time application of estimation intention loop 39

Figure 5.4.5: Elbow up intention in a real time application of estimation intention loop. 39

Figure 5.4.6: Intention of moving hand over head level in a real time application of estimation intention loop 40

Figure 5.5.1: Joint angle trajectories for Elbow Down 1 test from first set. 44

Figure 5.5.2: Joint angle trajectories for Elbow Up 1 test from first set. 44

Figure 5.5.3: Joint angle trajectories for Moving Hand Above Head 1 test from first set. 45

Figure 5.5.4: Probabilities of each intention for Elbow Down 1 test from first set. Blue trajectory is intention (1), red is intention (2), cyan is intention (3). 46

Figure 5.5.5: Probabilities of each intention for Elbow Up 1 test from first set. Blue trajectory is intention (1), red is intention (2), cyan is intention (3). 46

Figure 5.5.6: Probabilities of each intention for Moving Hand Above Head 1 test from first set. Blue trajectory is intention (1), red is intention (2), cyan is intention (3)........47

LIST OF TABLES

Table 4.2.1.1: Denavit Hartenberg Parameters 20

Table 4.2.2.1: Joint Rotation Effects 21

Table 4.2.3.1: Joint actuator specifications of arms 24

Table 5.2.1: This table gives the color representation of joint angle trajectories for all plots in the thesis. 30

Table 5.5.1: Results of Action Estimation Tests, Decision Cycle: 250 41

Table 5.5.2: Results of Action Estimation Tests, Decision Cycle: 100 42

Table 5.5.3: Results of Tests which were not goal oriented, Decision Cycle: 100........43

Chapter 1

1. INTRODUCTION

Fundamental questions regarding human consciousness have existed through ages and humans tried to answer these questions using the best set of explanations that was available at the time they lived. Contemporary explanations for consciousness both exist in philosophical branches such as philosophy of mind and in scientific branches such as computational theory of mind (CToM). Philosophy of mind is well outside the topic this thesis. On the other hand CToM is a branch in science where computer science and neuroscience intersects, and this thesis is based on certain findings on CToM on understanding actions of others.

Understanding actions of others is a subject that is being discussed in the neuroscience literature [1, 2, 3]. Although until early 1990s there was no empirical evidence in literature to connect action understanding to any part in human brain within the knowledge of the author. Thanks to his work on Macaque monkeys in 1980s and early 1990s G. Rizzolatti identified a certain set of cells called mirror neurons [4, 5].

After discovery of mirror neurons researchers such as V.S. Ramachandran tried to connect these neurons to action undertstanding mechanisms inside human brain [6]. Ramachandran even suggested that mirror neurons will advance the scientific work on brain in a way that discovery of DNA advanced the work on biology. There are also researches other than Ramachadan in literature that connects action understanding to mirror neurons [7, 8, 9]. Due to the recently discovered connection, researchers in the field of CToM were able to propose computational algorithms regarding action understanding based on mirror neurons.

Further on, researchers in the field of computer science are trying to explain results in brain research through computational models. Subtopic of finding computational models for brain parts is called CToM. [10, 11, 12] report computational models that can be used to understand human actions. [10] also proposes a computational model for understanding human actions. Moreover [10] connects its findings to mirror neurons. Most important quality of [10] for the context of this thesis is that proposed computational model is tested on robotic simulations. Therefore this work establishes a common ground for robotics and neuroscience.

Independently from developments in the field of neuroscience, the field of robotics has improved due to advancements in computer science. Contemporary robots have the possibility of electronically controlling its joints and visually processing data at the same time due to increased processing speed, and parallel computation. Considering these advancements it is possible to verify findings of CToM in a multidisciplinary framework through robotics. To this end [13] tries to define a framework which can be used to model robots which have as much agility as humans. Robotic actuation, sensing, and control mechanisms are described with respect to human muscles and human-like sensing and control terms.

Main motivation of the implementation in this thesis is to create a humanoid robot system which can estimate and even anticipate human actions. Although [10] proposes a computational algorithm, which has neurological basis, for estimating actions of others, the algorithm is not implemented on a robotic platform. The work in this thesis is on application of the proposed computational model on an actual humanoid robot platform. Another motivation is to show that robotic platforms can be used to simulate actual brain functions in humans to verify neurological findings. The implementation is also expected to make contributions to human machine interaction research by testing an algorithm which has the potential of being utilized in settings where humans and machines could work together.

Thesis is organized as follows. A survey on Human Robot Interaction (HRI) is given in the next chapter. Chapter three explains the neurological and technical ideas behind implementation of a computer mode in detail. A mental state inference/intention estimation loop based a forward model (FM) which was developed in [10] is explained along with its neurological basis. Chapter four explains the hardware used in implementation. Kinematic arrangement of arms of humanoid robot SURALP is presented along with Kinect, the visual input device used in the implementation. In chapter five details of the implementation of FM on SURALP are given. Moreover the test settings and results are given. All tests are made with FM proposed by [10]. Eventually chapter six concludes the thesis by discussing results of tests and limitations of the computational model, and by presenting the future work.

Chapter 2

2. A Survey on Human Robot Interaction

HRI is a new topic that emerged in the late 20^{th} century. It is a multidisciplinary topic which combines artificial intelligence with robotics to create robots which can respond to humans. HRI is a vast topic that has applications on different kinds of robots. Although fully autonomous industrial robots have some sort of interaction with their human operators/programmer, HRI research is based on more complicated interaction scenarios. These scenarios between robots and humans are possible due to utilization of low level robot control algorithms to create goal oriented physical actions which have sophisticated meanings.

Earliest research in the field can be traced backed to beginning of 1990s [14, 15]. [16] tries to explain intelligence in terms of combination of low and high level sensory feedback loops. [16] also makes comments on possible applications of its explanations on robotic platforms. [17] is one of the earliest works that explains action anticipating capabilities of robotic systems. It also proposes that these capabilities can already be observed in nature.

Some important domains in human life for HRI are rescue operations, developing robots with medical applications for children with autism, and possible applications of brain research on robotic platforms to create human-like robots [18, 19, 20, 21, 22, 23]. [24] proposes a computational model for robot that allow them to follow humans. Such an application can be utilized in both in medical setting where patients require assistance of robot and in rescue scenarios.

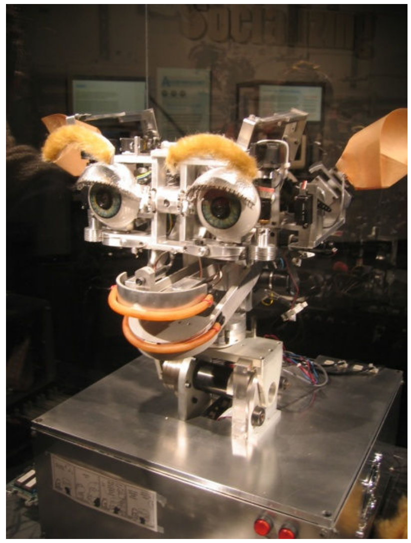

Figure 2.1: Kismet, a robot which is capable of interacting with humans through understanding and mimicking emotions

Over the last 20 years many robots with capabilities which allow them to interact with humans have emerged. Kismet, shown in Figure 2.1., is a robot that was developed in Massachusetts Institute of Technology which can mimic human emotions. SONY's AIBO is a commercial pet robot in the shape of a dog [25], it has the ability to interact with kids mainly for entertainment. Figure 2.2 shows SONY's AIBO. Philos is another example, it can interact with humans to monitor their health status [26]. Philos is shown in Figure 2.3.

Figure 2.2: SONY AIBO ERS7A

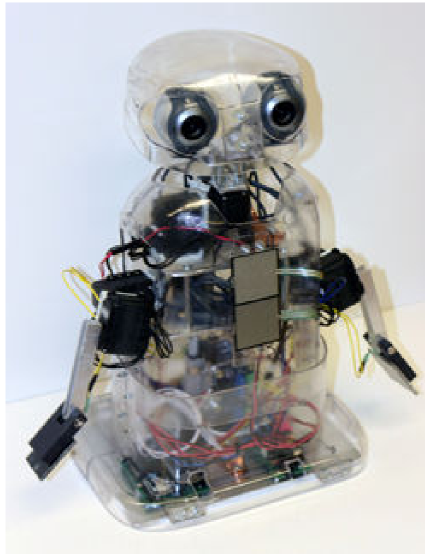

Figure 2.3: Philos, a social robot developed at Case Western Reserve University

Boston Dynamics is one of the robotic design companies which leads the research on robotics. Although founded in company was founded in 1992, they are best known for Big Dog, a quadruped designed for operating in unstable terrains [27]. Atlas is very new robot which was developed by Boston Dynamics, introduced in July 11, 2013. It is distributed to research institutions for developing artificial intelligence systems which will be capable of making decisions in environments which are dangerous for humans to work. It is also planned that Atlas will be working with humans on rescue operations. Atlas is shown in Figure 2.4.

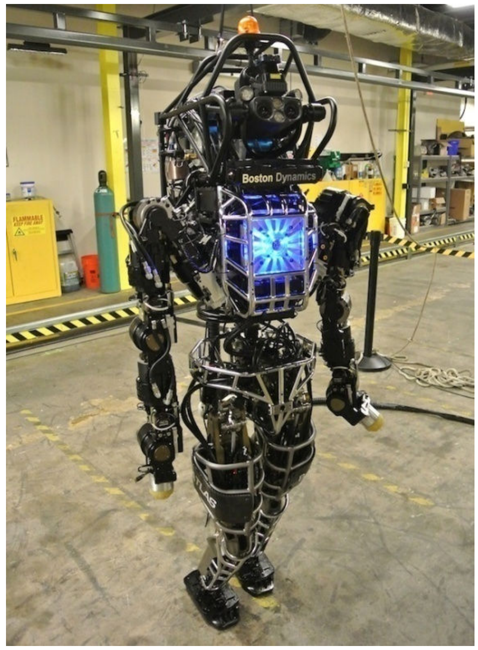

Figure 2.4: Atlas, designed and produced by Boston Dynamics

Research based on robotic exoskeletons is another topic related to HMI. An exoskeleton must have stable dynamics, and it should be able to remove excessive forces which can injure human operators. Exoskeleton is in constant interaction with the human operator in order to track his/her movements. BLEEX is an exoskeleton system developed by H. Kazerooni. It can be seen in Figure 2.5. BLEEX has functions which can increase human operators' physical capacities. [28] is another research based on an exoskeleton which increases physical capacity during walking. There are ongoing researches on medical applications of robotic exoskeletons [29, 30, 31, 32].

Figure 2.5: BLEEX was developed by Berkeley Robotics and Human Engineering Laboratory

Anticipating actions of humans and moving in a way which can assist them is also one of the domains of HRI. This is also the part of HRI research that is in the scope of this thesis. Though, not much research is made in the context of action anticipation, a very recent study reports a method for action anticipation which calculates costs of reaching certain objects [33]. Unlikely path scenarios are eliminated after costs are calculated, and robot executes movement after anticipating the actor's action. Another study reports training of robots from human movements for inferring intentions [34].

Chapter 3

3. Basis of Thesis

First section of this chapter explains the neurological ideas behind a technical realization of a computational model which estimates human intentions. The neurological topics in the first section are Theory of Mind (ToM), Mirror Neurons and brain research literature which reports findings on how humans and animals understand actions of others.

In the second section, reasons of a technical realization are discussed. The computational model behind such realization is briefly explained, a more detailed explanation is given in Chapter 4. Eventually the setting in which a robot could estimate human intentions is described.

3.1. Neurological Basis of Intention Estimation

Humans and animals automatically detect movement and come to conclusions which are going to help them survive. Moreover humans detect movements and make inferences on others mental states from these movements for social interaction. Therefore understanding intentions of others from observed data can be regarded a subtopic of ToM. According to [35] a ToM is the ability to realize that others have beliefs and intentions other than us, and the ability to come to conclusions on their behavior with respect to their beliefs and intentions.

Human brain is capable of processing various inputs. Visual sensory system in the brain is one of these processing capabilities. Visual data obtained from sensory system that can be visual patterns such as face motions, and change of visual patterns with respect to time or body limb positions and their change with respect to time. There is evidence in literature that humans utilize obtained visual data in ToM models inside the brain to simulate beliefs on others intentions [36, 37, 38].

It is also argued that ToM is not only applicable to humans. [39] states that both humans and chimpanzees' have some sort of ToM. Both species have mechanism that understands observed goals of other humans, but chimpanzees' lack a mechanism which detects false beliefs. A false belief is the information that a trusted agent has the wrong answer to a question and it is a critical point in determining the complexity of ToM models inside brains of different species. A test on false beliefs is conducted in [40]. Results show that ToM of chimpanzees' is not able to detect false beliefs while infants of age 4-5 are successful in detecting them. These results point out that ToM system is much simpler even in animals which are closely related to humans, yet it is able to produce meaningful results regarding others intentions.

Another research area that builds the neurological basis is the computational ToM. This was proposed by Hilary Putnam in 1961. It basically states that human brain is an information processing device and its activities can be explained by computational models. Steven Pinker proposes that intentions are stored in mind as information and information is then used to create sophisticated decisions [41]. If computational theory mind is taken seriously it can be concluded that our physical actions are result of layers of computation made by brain.

Mirror neurons are a recent finding in the field of neuroscience. They are believed to be related to activities that are attributed to imitation and learning in the brain. In 1980s and early 1990s Giacomo Rizzolatti and his research team came upon the evidence of mirror neurons while experimenting on Macaque monkeys. According to [4] a mirror neuron is a cell which can fire both when an action is observed and when an action is performed. Neuroscientists are working on mirror neurons to show that these neurons play an important role in producing intentions by simulating observed actions of an actor [42, 43, 44, 45]. Researchers have empirical data from humans and animals that an internal ToM is actually existent in brains. Mirror neurons can provide consistent ways for computationally modeling and verifying findings of research based on ToM, especially in the field of action understanding [46, 47].

[48] reports an experiment in which motions of the bodies of human subjects were represented with light sources in a dark room. Observing humans were able to gather efficient knowledge from light sources when the number was 10-12, and observers were able detect walking patterns when number of sources were as low as 5. If research based on ToM is assumed to be accurate, it could be concluded that there is a model in the brain which can work with very small amount of visual data to infer the action conducted by the actor.

Oztop, Wolpert and Kawato discusses in their work that humans have a FM which helps them to estimate/understand observed actions [10]. They also argue that the intention estimation loop in the brain and the proposed FM are consistent with the finding on mirror neurons and findings on action estimation theories based on ToM. Their proposed FM simulates an actor's movement and then an intention estimation loop compares results with the actual observed data. More neurological and computational details on the FM and the action estimation loop are given in the next section. Chapter 5 explains the FM and action estimation algorithm in full detail.

3.2. Technical Realization of Intention Estimation

A technical realization of a model which can estimate/understand human intentions have practical applications. It can be utilized to test findings of brain research, the findings of action understanding. At some point in the future technical realization can be used to pave the ground for more complex models which can be used to treat people with mental difficulties. It can also have practical applications in human-machine interactions. Robots who think like humans can also work with humans in different settings [49, 50, 51].

[10] provides a computational algorithm to estimate human intentions which is also consistent with findings in brain research literature. It is stated that a FM, which is composed of mirror neurons, is inherited in human brain. It is also stated that this model has two purposes: (1) Anticipating visual consequences of execution of a goal oriented action in order to compensate for visual feedback delay (2) simulating mental states of others using the observed visual data and sending it to an intention estimation loop in order to come to conclusions on their actions. Although both of these purposes can be implemented on robotic platforms, scope of the thesis covers application of second purpose on a robotic platform. A robot platform which has access to such a FM and intention estimation loop would be able to detect human motivations.

In the most basic setting, a robot arm, a visual sensor and a computer which process the visual data and passes it to the FM algorithm are required for such a technical realization. The computer and the robot should also be able to communicate with each other. Robot arm must have a similar workspace to that of human arm. Therefore kinematic arrangement of the robot arm should be anthropomorphic. Its parameters should be clearly defined, so that visual data obtained from humans can be matched with robot control parameters.

A discussion can be made on the location of the FM in a technical realization can be made at this point. [10] states that exact location of FM is uncertain. It could either be located in the part of premotor cortex of humans, which is believed to have an effect on planning of movement that contains mirror neurons, or it could be located as a combination of models distributed over cerebellum, part of brain which plays an important role in motor control, and premotor cortex. Since exact location of the FM is ambiguous, in the technical setting it can be regarded as an executable running in the computer which receives visual data. Same executable is also capable of running an intention estimation loop.

Figure 3.2.1 shows the mentioned brain parts of the last paragraph.

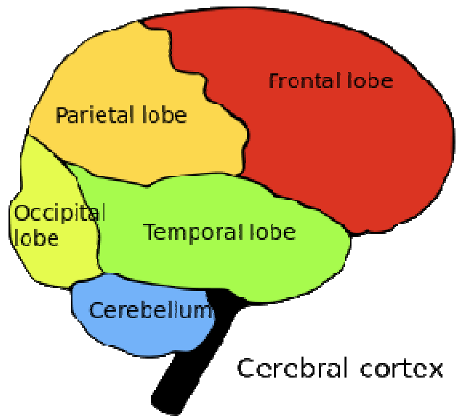

Figure 3.2.1: Cerebral cortex of human brain is shown. It is the outermost layer of neurons in the brain. Premotor cortex is located in frontal lobe. Cerebellum can be seen in the bottom.

Chapter 4

4. Hardware Used in Thesis

This chapter explains the details of hardware that were used in research. These hardware choices are consistent with the proposed technical realization in the previous section.

First section describes the visual input sensor and explains why it was chosen. Second section explains physical and kinematic properties of the humanoid robot platform. Terminologies regarding modern robotics are also emitted in this section.

4.1. Kinect

Kinect is a visual input device developed by Microsoft, which has motion sensing capabilities. For the scope of the work in the thesis, it is used to recover transformation matrices related to hand, wrist, elbow and shoulder. Camera frame convention for Kinect is shown in Figure 4.1.1. Camera frame is denoted as f_c in this work. All transformation matrices obtained from Kinect are represented with respect the coordinate system in the figure.

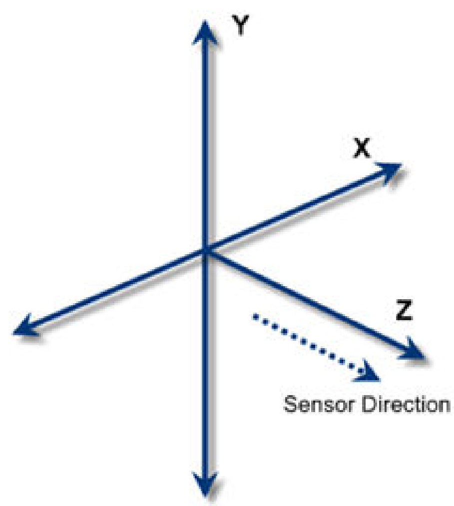

Figure 4.1.1: In the coordinate system: z represents the direction which camera is looking at, x is the direction to the left of the device, and y is perpendicular to x and z.

Figure 4.1.2 shows the human skeleton can be captured by Kinect. It is possible to get Cartesian position information of certain body parts. Moreover Kinect has the option to compute locations of body parts in the shoulder frame of the actor. Shoulder frame coordinates for Kinect are given in Figure 4.1.3. . Shoulder frame for Kinect skeleton is denoted as $f_{shoulder_{camera}}$ in this thesis.

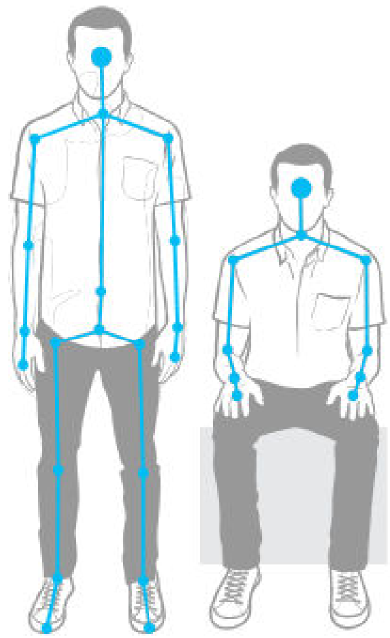

Figure 4.1.2: Two different skeletons captured by kinect are shown. Left is the skeleton of a standing person, right is the seated version.

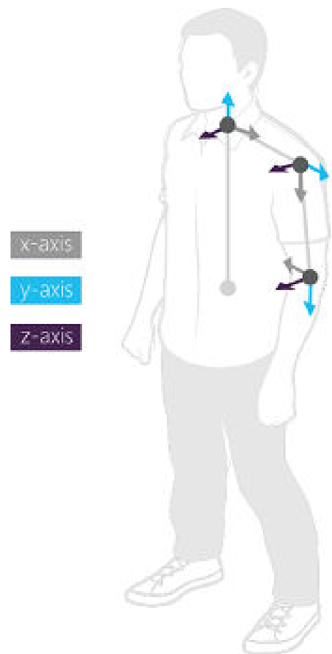

Figure 4.1.3: Coordinate axis assignments of Kinect Shoulder frame is given along with neck and elbow.

One particular task that Kinect is able to do is to compute the transformation matrix of human shoulder. This problem is determining 3D transformation of an object/a body part from a 2D image. In the modern 3D vision literature the problem is addressed as the pose estimation problem. One of the most common pose estimation algorithms is POSIT [52]. It uses a known 3D model of an object to compute the 3D transformation matrix from image coordinates. Other solutions to pose estimation problem are also discussed in the literature [53, 54]. Details on how Kinect solves this problem are explained in [55]. The most important reason to use Kinect to gather visual inputs is that Kinect has a built in library that solves pose estimation problem.

4.2. SURALP: A Full Body Humanoid Robot

SURALP is a humanoid robot at Sabanci University. The project was funded by TUBITAK, and it was originally developed to conduct walking experiments. It can track circular trajectories and enter sloped surfaces [56, 57]. Figure 4.2.1 shows SURALP's current physical appearance.

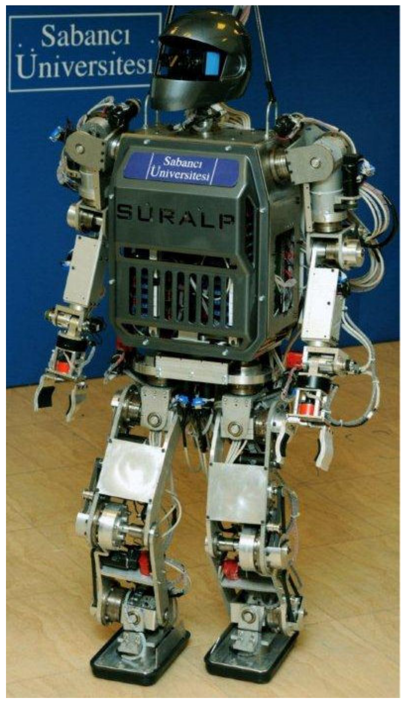

Figure 4.2.1:SURALP

SURALP's kinematic arrangement consists of 29 Degrees of Freedom (DoF). They are distributed as follows: 2 at neck, 1 at hip, 6 at each leg, 6 at each arm, 1 at each hand. The DoF at hand is a gripper which produces linear motion for grasping objects. SURALP is 166cm long and weighs 114 kg. Dimensions of SURALP are given in Figure 4.2.2. SURALP has CCD cameras for visual data, but those cameras are not used in this work. Instead a Kinect camera is used as explained in the previous section.

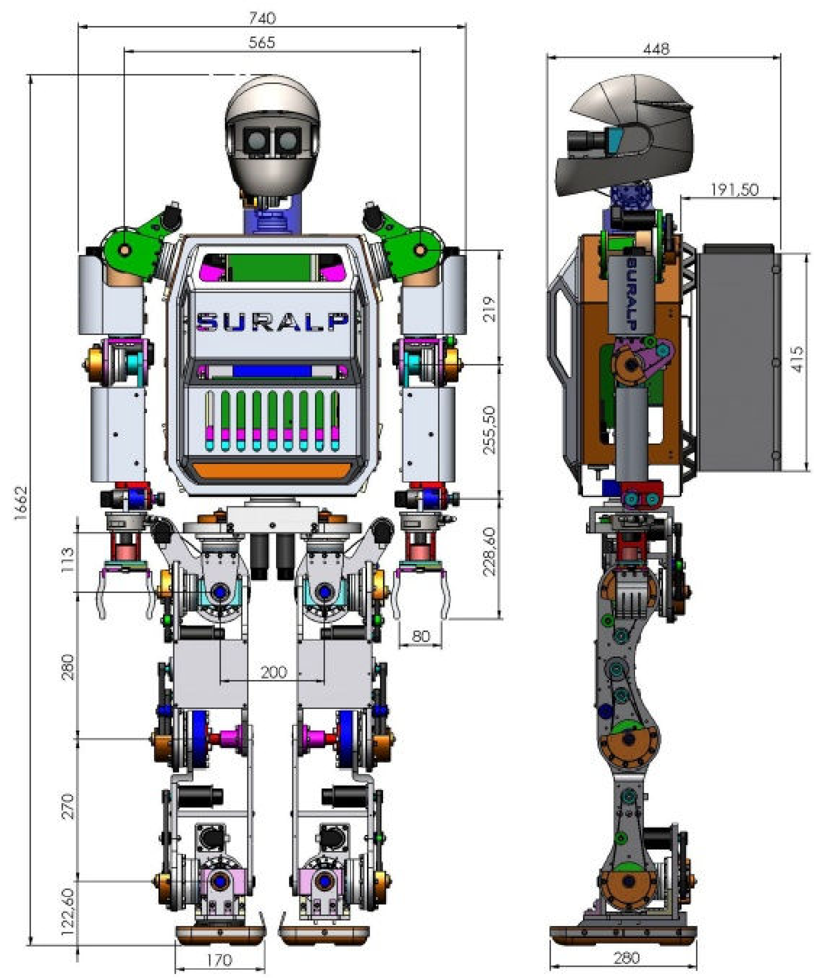

Figure 4.2.2: Dimensions of SURALP in milimeters

4.2.1. Kinematic Arrengement of Arms

Design of SURALP's arm is anthropomorphic. It is composed of 3 rigid bodies: upper arm, lower arm and a hand. It has a total of 6 DoF: 3 DoF at shoulder, 2 DoF at elbow, 1 DoF at wrist. A human arm is considered to have 2 DoF at wrist therefore making a total of 7 DoF. This distinction does not create inconveniences for the scope of the research because 6 DoF are enough to control all physical DoF, 3 positions and 3 orientations. DoF in arms can be seen in Figure 4.2.1.1.

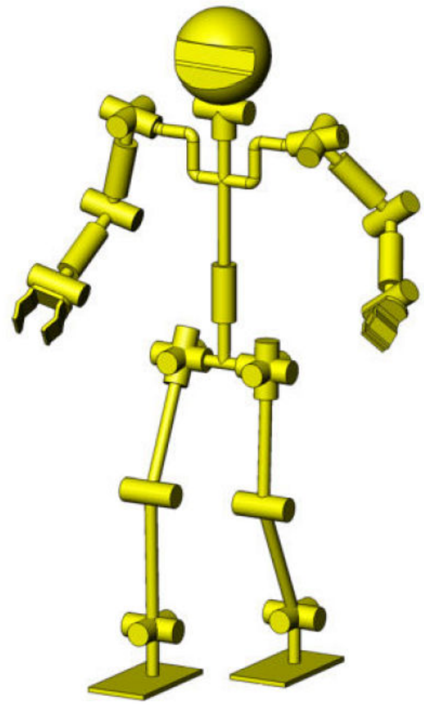

Figure 4.2.1.1: Kinematic Arrengement

One of the main problems while working with multi DoF robots is to map motor joint angles to Cartesian coordinates and orientation of robot end effectors. This is called the forward kinematics problem. Obtaining joint angles from Cartesian position and orientations is called inverse kinematics problem. This section explains how forward kinematics problem is solved while working on SURALP.

SURALP has a coordinate system at the center of its trunk, f_{trunk}. x-axis of the coordinate system is along the walking direction, z is up and y is pointing left. SURALP has another coordinate system at the base of its arm. The second coordinate system will be denoted as base shoulder frame, $f_{shoulder}$. It differs from the coordinate system at the center of its trunk by a 15 degree rotation along x-axis of f_{trunk}.

Denavit-Hartenberg parameterization is used in order to compute the transformation matrices which relate $f_{shoulder}$ to frame that is located at the hand, f_{hand} . Denavit-Hartenberg parameters for each arm are shown in Table 4.2.1.1, and axis assignment are shown in Figure 4.2.1.2. The transformation matrix which relates $f_{shoulder}$ to f_{hand} contains information regarding Cartesian position and orientation of hand with respect , $f_{shoulder}$. A content of such transformation matrix is shown in (4.2.1.1).

$$T_{f_{hand}}^{f_{shoulder}} \begin{bmatrix} 3x3\ orientation\ matrix & 3x1\ Position\ vector \\ 0 & 1 \end{bmatrix} \tag{4.2.1.1}$$

In equation .. $3x3\ orientation\ matrix$ defines orientation of f_{hand} with respect to $f_{shoulder}$ and $3x1\ Position\ vector$ defines Cartesian position of the origin of f_{hand} in $f_{shoulder}$. $T_{f_{hand}}^{f_{shoulder}}$ is obtained by multiplying transformation matrix of each joint obtained from Denavit-Hartenberg convention. Forward kinematics problem is solved by computing the $T_{f_{hand}}^{f_{shoulder}}$. Solution to inverse kinematics problem is explained in the next subsection.

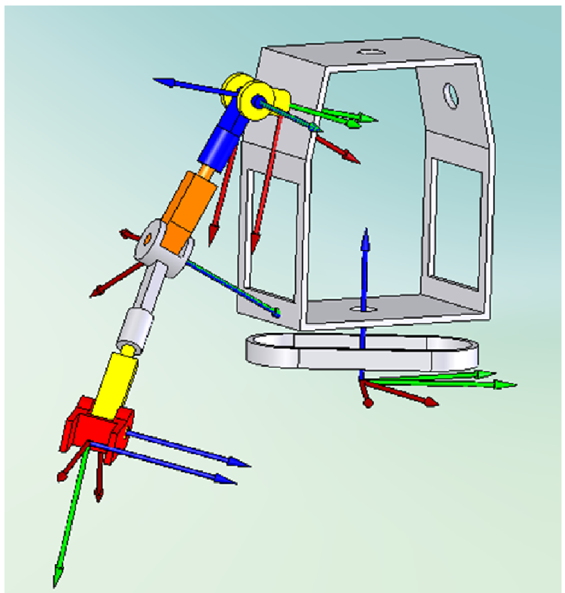

Figure 4.2.1.2: Denavit-Hartenberg based axis assignment for 6-DoF arm. Green arrows denote y axes, red x axes and blue z axes.

Table 4.2.1.1: Denavit Hartenberg Parameters

	a	α	d	θ
Link 1	0	-90°	0	θ_1^*
Link 2	0	90°	0	θ_2^*
Link 3	L_3	0°	0	θ_3^*
Link 4	L_4	0°	0	θ_4^*
Link 5	0	-90°	0	θ_5^*
Link 6	L_6	0°	0	θ_6^*

4.2.2. Inverse Kinematics Based on Visual Data Obtained from Kinect

As briefly explained in the previous subsection inverse kinematics is the problem of obtaining joint angles of a multi-dimensional system from Cartesian positions and orientation information of the end effector. Usages of inverse trigonometric functions make the inverse kinematic problem a highly non-linear one.

Inverse kinematics solutions of SURALP based on hand positions and orientations are well defined, but these solutions are omitted in this thesis. This is due to the fact that hand orientation data obtained from Kinect is too noisy to be worked with. Instead of using hand positions and orientations to solve inverse kinematics; shoulder orientation, elbow, wrist and hand positions are used.

When Denavit-Hartenberg parameters shown in Table 4.2.1.1 and axis configuration in Figure 4.2.1.2 are used with a 6-DoF robot arm, 6 joint angles can be obtained from elbow, wrist and hand positions with the formulation shown in equations (4.2.2.1) through (4.2.2.6). In equations superscripts denote the frame in which x and y positions are expressed. Transformation matrices are used to define positions in different frames.

$$\theta_1 = atan2(y_{elbow}^0, x_{elbow}^0) \tag{4.2.2.1}$$

$$\theta_2 = atan2(y_{elbow}^1, x_{elbow}^1) + {}^{pi}/_{2} \tag{4.2.2.2}$$

$$\theta_3 = atan2(y^2_{wrist}, x^2_{wrist}) - {}^{pi}/_{2} \tag{4.2.2.3}$$

$$\theta_4 = atan2\left(-\sqrt{1-\rho^2}, \rho\right) \tag{4.2.2.4}$$

$$\theta_5 = atan2(y^4_{hand}, x^4_{hand}) + pi \tag{4.2.2.5}$$

$$\theta_6 = atan2\left(y^5_{hand}, x^5_{hand}\right) \tag{4.2.2.6}$$

In (4.2.2.7) ρ is cosine of θ_4. (4.2.2.7) is obtained from cosine theorem between two links of the arm, upper arm and lower arm and it shows computation of ρ along with (4.2.2.8). Upper arm is the link between shoulder and elbow, and lower arm is the link between elbow and wrist. These links can be observed in Figures 4.2.2, 4.2.1.1 and 4.2.1.2.

$$c = \sqrt{(x^0_{wrist})^2 + (y^0_{wrist})^2 + (z^0_{wrist})^2} \tag{4.2.2.7}$$

$$\rho = \frac{c^2 - a^2 - b^2}{2ab} \tag{4.2.2.8}$$

In (4.2.2.8), a denotes length of upper arm, and b denotes length of lower arm.

It should be noted that 0^{th} frame of the arm is regarded as base shoulder frame, $f_{shoulder}$. On the other hand visual data obtained from Kinect uses f_{camera}. The formulation in (4.2.2.9) must be used to convert a point P expressed in f_{camera} to $f_{shoulder}$ using a constant orientation matrix $R^{f_{shoulder}}_{f_{camera}}$.

$$P^{f_{shoulder}} = R^{f_{shoulder}}_{f_{camera}} P^{f_{camera}} \tag{4.2.2.9}$$

Where;

$$R^{f_{shoulder}}_{f_{camera}} = \begin{bmatrix} -0.7071 & 0 & 0.7071 \\ 0 & 1 & 0 \\ -0.7071 & 0 & 0.7071 \end{bmatrix} \tag{4.2.2.10}$$

Equation .. indicates that f_{camera} and $f_{shoulder}$ can be aligned by a rotation of ${}^{pi}/_{4}$radians in y-axis of f_{camera}. f_{camera} and $f_{shoulder}$ can be seen in Figures 4.1.2 and 4.1.3 respectively.

Visual data obtained from Kinect are in the following form:

$$T_{f_{shoulder_{camera}}}^{f_{camera}} = \begin{bmatrix} R_{f_{shoulder_{camera}}}^{f_{camera}} & 3x1\ Position\ vector \\ 0 & 1 \end{bmatrix} \tag{4.2.2.11}$$

$$P_{hand}^{f_{shoulder_{camera}}} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{4.2.2.12}$$

$$P_{wrist}^{f_{shoulder_{camera}}} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{4.2.2.13}$$

$$P_{elbow}^{f_{shoulder_{camera}}} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{4.2.2.14}$$

From equations (4.2.2.11) through (4.2.2.14) positions of hand, wrist and elbow are obtained, but these positions are in $f_{shoulder_{camera}}$. From the trasnformation matrix in (4.2.2.11) $R_{f_{shoulder_{camera}}}^{f_{camera}}$ can be obtained. Now if (4.2.2.15) is combined with the rotation matrix in (4.2.2.11). each position vector defined in (4.2.2.12), (4.2.2.13), (4.2.2.14) can be expressed in $f_{shoulder}$ using the following equation.

$$P_{hand}^{f_{shoulder}} = R_{f_{camera}}^{f_{shoulder}} R_{f_{shoulder_{camera}}}^{f_{camera}} P_{hand}^{f_{shoulder_{camera}}} \tag{4.2.2.15}$$

Only the formulation for hand position is shown in (4.2.2.15), formulation of wrist and elbow are similar. Results of (4.2.2.15) can be used in the proposed inverse kinematics formulation since positions are expressed in the 0^{th} frame, i.e. $f_{shoulder}$.

4.2.3. Actuation of Arms

There is a DC motor at each joint of arm. Joints are able to track position references which are either obtained from inverse kinematics or directly applied by users. The electronic hardware which controls the joints is the dSpace control desk. Central control board is DS-1005. A DS 3001 board is connected to encoders which are located at each motor drive that read current position references and use PID controllers to drive them to reference positions. DS 2002 board is used to convert analog data obtained from force/torque sensors to digital, and a DS2103 board is used to convert digital reference signal data to analog data which is to be sent to actuators. Figure 4.2.3.1 shows the mentioned electronic control boards in an hierarchical setting. Table 4.2.3.1 shows motor powers, and ranges of arms.

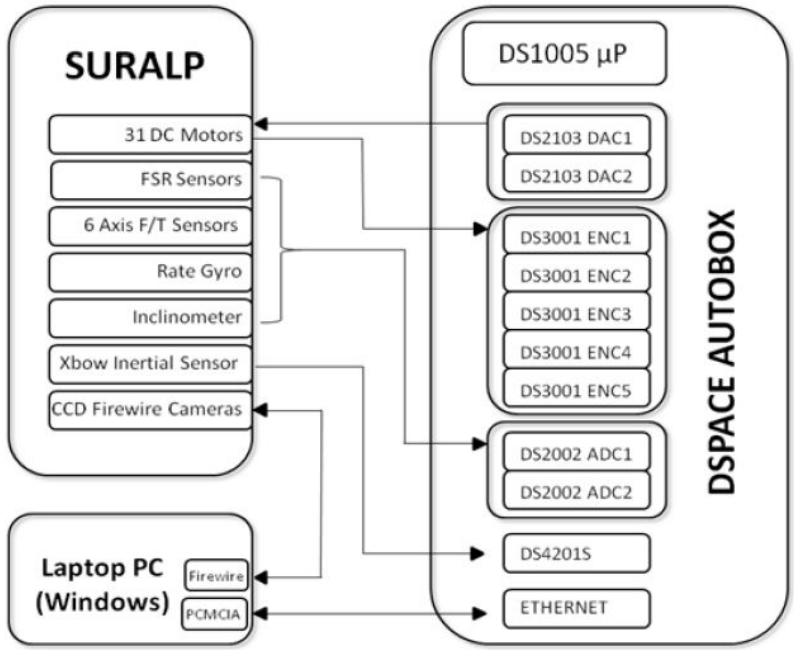

Figure 4.2.3.1: Hardware Architecture of SURALP

Table 4.2.3.1: Joint actuator specifications of arms

Joint	Motor Power	Motor Range
Shoulder Roll 1	150W	-180 to 180 deg
Shoulder Pitch	150W	-23 to 135 deg
Shoulder Roll 2	90W	-180 to 180 deg
Elbow	150W	-49 to 110 deg
Wrist Roll	70W	-180 to 180 deg
Wrist Pitch	90W	-16 to 90 deg
Gripper	4W	0 to 80 mm

Chapter 5

5. A Humanoid Robot Platform Capable of Understanding Intentions of Others

First section of this chapter explains the dual function of the proposed FM in human brain. It also explains how FM could work in the actual human brain. Computational details of the FM are in the second section. The FM used in [10] and the FM in this thesis have differences. These differences are pointed out in Section 3.

Section 4 explains the settin in which all tests are made. Section 5 presents details on experiments conducted with the FM and their results.

5.1. Neurological Details of FM

As [10] states there is enough evidence in literature to safely assume that humans have models that make estimation on actions on others [37, 48, 58, 59]. [10] proposes a computational FM to explain the estimation process. The FM gets activated whenever one of two goal directed events is observed: (1) When a goal directed event is executed (2) When a goal directed event is observed from others. In both cases primary objective of the FM is to predict future intentions. It is stated that this model has dual advantages of reducing sensory delays while executing an action and estimating actions while observing others. The FM will be used in observation mode in for the scope of the thesis. Structure of model for the observation mode can be observed in Figure 5.1.1. (//) in the figure indicates the disconnected paths in the brain while FM is used for intention estimation. (//) are switched in movement execution scenario, therefore they become connected to brain parts which actually creates movements in muscles.

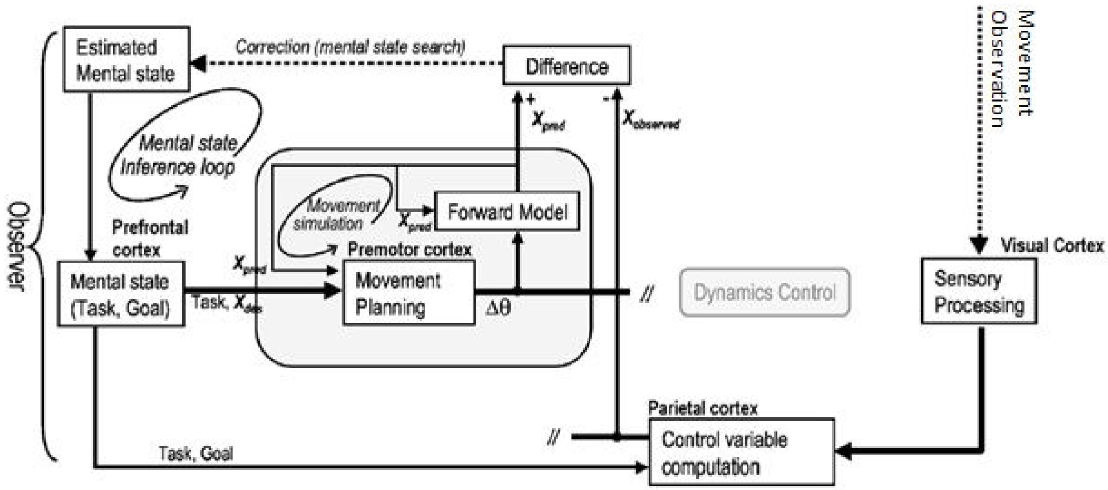

Figure 5.1.1: Block representation of the proposed FM for action estimation in human brain. Image is obtained from [008].

Intention estimation/mental state inference loop uses data obtained from two different parts of brain to make decisions. These parts are Parietal Cortex and the proposed FM. Parietal Cortex of human brain is believed to be responsible for gathering processing visual reaching data [60, 61, 62]. Visual data is then sent to intention estimation loop. The data is named as observed control vector $X_{observed}$. On the other hand the FM creates simulated control vector $X_{predicted}$. [63] states that Premotor Cortex is activated while human brain is selecting a movement from a set. The FM receives parameters of a movement from the Premotor Cortex. This data is used to create $X_{predicted}$. Eventually $X_{predicted}$ is also sent to intention estimation loop to come to a decision regarding movement of the actor..

ToM states that humans understand other human's beliefs and actions by simulating them inside their brains. From context of ToM the proposed intention estimation mechanism along with the FM is able to simulate other people intentions to understand/estimate their actions.

According to [10] there are unconscious brain activities during action observation of actions in the parts that are believed to contain large amounts of mirror neurons. It is proposed that simulation of the proposed FM during action observation can explain the activity. Therefore one neurological explanation for such a mental simulation scenario is activation of mirror neurons during observation.

5.2. Computational Details of FM

Two similar computational models are proposed in [10]. One of these models makes search in a finite set which is composed of different intentions/mental states. A discrete set of actions, such as reaching an object with elbow down or elbow up configurations, can be searched using the first model. The other model is capable of making decision in sets with infinite elements. Estimating intention of an actor while s/he is reaching a certain point in a plane can be given as an example to applications of second model. The sets of points on the plane can be regarded as a set with infinite points. Only the computational algorithm for the first model is implemented in the robotic arm.

In the last section $X_{observed}$ and $X_{predicted}$ were defined while explaining the intention estimation mechanism inside human brain. In this section $X_{observed}$ is denoted as X_o and $X_{predicted}$ is denoted as X_e. The reason for the change of expression is that first notation is adapted by [10] and details of the FM in [10] are slightly different than the FM used in the thesis. These differences are explained in detail in Section 5.3.

The computational algorithm for such a FM is proposed in [10] but it is not implemented in an actual robotic platform. In order to implement the model, X_o and X_e are defined as $5x1$ column vectors which respectively represent joint angles of the actor with respect to kinematic arrangement of robot and estimated joint angles of the robot. Computation of X_o and X_e are explained in the following paragraphs.

Although the robot arm used in experiments have 6 joints, control vectors are composed of 5 joint positions. Arctangent function inside inverse kinematics formulation occasionally produces answers outside $\{-2pi, 2pi\}$ range for 5^{th}joint angle of the robotic arm from observed Cartesian arm positions, resulting in discrete jumps in data. Another problem with 5^{th}joint even in cases with no discrete jumps, it is still too noisy to be worked with. In order to obtain non-noisy differences between control vectors, 5^{th}joint angle is omitted. Time trajectory of 5^{th}joint angle can be observed in Figure 5.2.1.

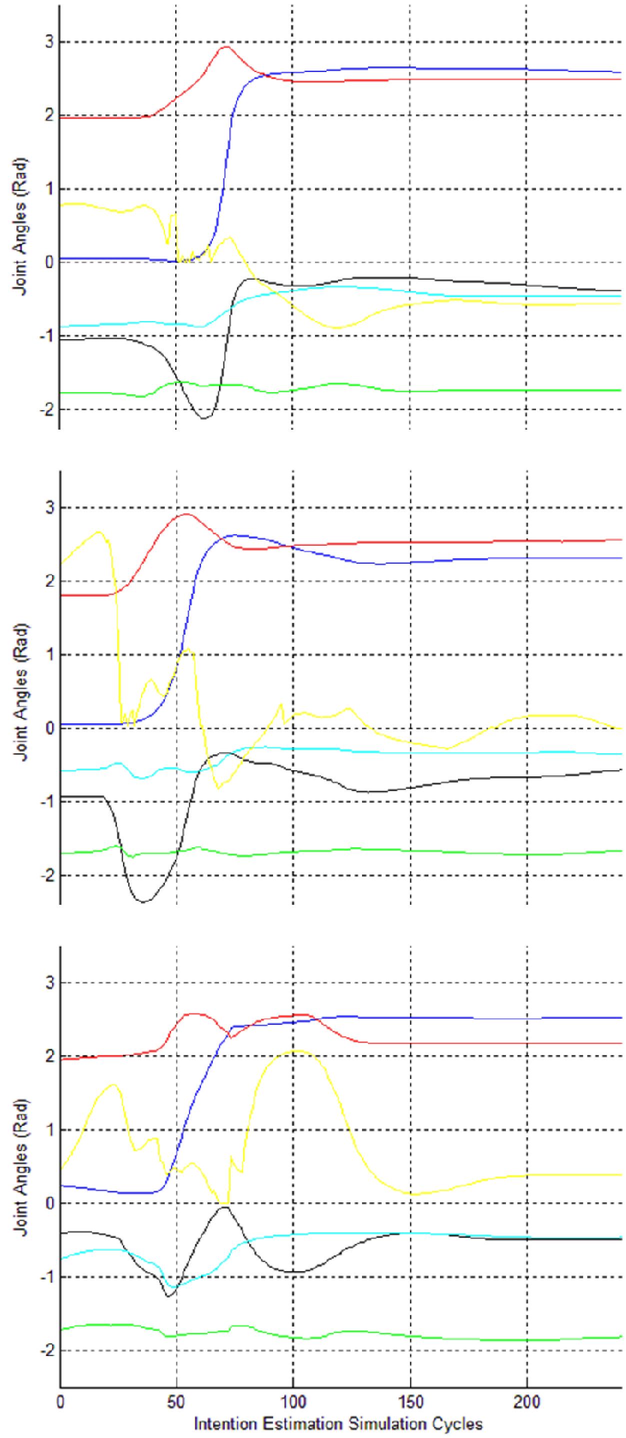

Figure 5.2.1: Three different trajectories for the intention of moving hand above head level is shown. Each joint trajectory is represented by a color. Color representations are given in Table 5.2.1.

Table 5.2.1: This table gives the color representation of joint angle trajectories for all plots in the thesis.

Joint Number	Color
1	Blue
2	Red
3	Black
4	Cyan
5	Yellow
6	Green

Trajectories of all joint angles have similar patterns for the same intention. The intention is to move the hand above head level. In Figure 5.2.1 three shoulder joints are activated in the beginning of the movement. These shoulder joints angle trajectories are shown in blue, red and black. Physical results of activation of these joints are discussed in Table 4.2.3.1. This certain behavior marks the intention of moving hand over head level in terms of SURALP's kinematic arrangement. On the other hand 5^{th}joint angle trajectories, shown in yellow in Figure 5.2.1 vary with unexpected behavior.

After explaining contents of X_o and X_e, details of the intention estimation algorithm used in the robotic arm can be given as follows:

1- Detect the beginning of goal oriented action
2- Initialize T_o and T_e to empty matrices
3- Compute the following at each cycle until a decision is made:
   a) Observe X_o and store it in T_o
   b) Simulate X_e until length of T_o for all possible intentions and store in T_e
   c) Calculate the difference D between each T_e and T_o from (5.2.1)
   d) Find the smallest D among all intentions
4- Stop the algorithm if a decision cycle is reached
5- If $D$is larger than a certain threshold do not make a decision

$$D^N = \frac{(1-\gamma)}{(1-\gamma^{N+1})}\sum_{i=0}^{N}(T_e(i)-T_o(i))^T \mathbf{W}(T_e(i)-T_o(i))\gamma^{N-i} \tag{5.2.1}$$

In (5.2.1) $T_e(i)$ and $T_o(i)$ denote i^{th}indices of T_e and T_o matrices. These matrices initially have unknown length and are filled with control vectors until the end of simulation. Their terminal sizes are $5xN$, where Ndenotes the last cycle of the simulation. γ is a constant real number which allows equation to put more weight on latter entries in T_e and T_o . γ was picked as 0.9 in all simulations. $\mathbf{W}$ is a diagonal $5x5$ matrix. Diagonal entries are picked as 0.2 in all simulations.

Detecting beginning of an observed movement is crucial in a real time application of FM algorithm. Difference between current and previous X_o are computed at each cycle. Entries of difference vector are stored in a moving matrix of length 20. The term "moving" is used because the matrix stores the difference vector from current cycle to 19 cycles before. Entries of resulting moving matrix are summed, if result is higher than 0.05, FM infers that a movement by the actor is initiated. The 0.05 threshold is determined prior to execution of FM model algorithm from differentiated X_o values. In Figure 5.2.2 time differentiation of X_o values are shown. These plots played an important role while choosing an appropriate threshold value for detecting movement.

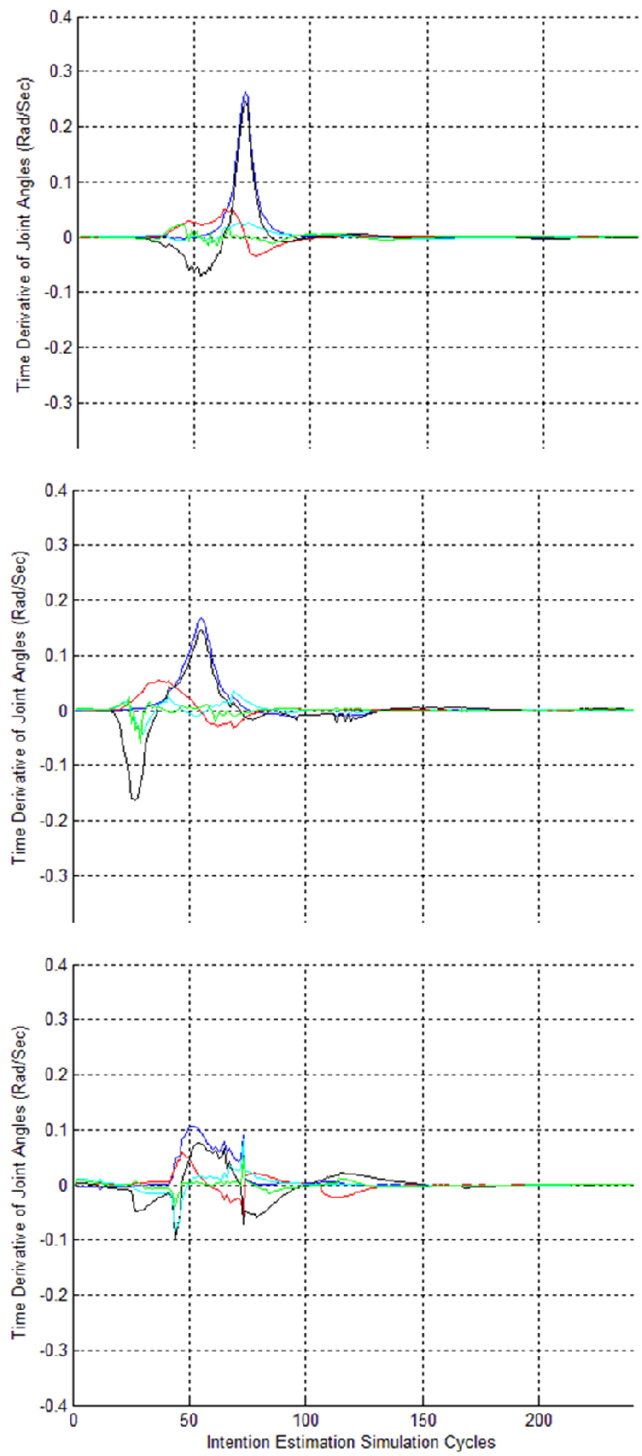

Figure 5.2.2: These plots show time differentitaion of joint angle trajectories of three different moves from Figure 5.2.1. Color representations are in Table 5.2.1. Notice that 5^{th}joint angle trajectories are omitted in time differentiations.

In order to compute the difference between movements of an actor and robot, a common framework must be established. In the thesis this framework is the joint space of the robot. Each visual data regarding positions of actor's arm are collected at each cycle and converted to joint angles of SURALP. The form of visual data and its conversion to joint angles of SURALP are explained in Chapter 4. These angles are then matched with SURALP's own time trajectory of joint angles using (5.2.1). SURALP's own time trajectories regarding different intentions are computed off-line. Computational details of time trajectories of different intentions are given in the next section.

5.3. Modifications of Thesis to Proposed FM

There are three modifications that are made in computational FM algorithm to utilize robot platform to estimate the intention of observed actor. The first modification is related to movement detection of FM. Latter is related to the decision cycle. Last one is related to X_e, the estimated control vectors.

At this point it should be noted that the method used in previous section to detect movements, detects all kinds of arm movements, whether or not a movement is goal oriented. Therefore D values are compared with a threshold at the decision cycle to eliminate movements which are not goal oriented. It should also be noted that this elimination is not based on [10] and it does not have a neurological basis.

Decision cycle, which was defined in the algorithm in the previous section, is the cycle in which FM in robot comes to a conclusion on actor's intention and starts to assist the actor. This decision cycle is the second modification to the proposed FM. In the [10] version of the FM algorithm there is no particular decision cycle and simulated robot do not move to assist the actor. Although it is proposed that observer/robot will have an accurate estimate on intention of actor until halfway through the movement. Decision values in experiments are picked with respect to this acknowledgment.

In modified FM algorithm X_e are computed from previously observed hand, wrist, elbow and shoulder coordinates. These coordinates are then converted to joint angles and stored in matrices to be used to calculate the difference between T_o and T_e. Therefore they do not have dynamics of their own; they do not have different values for different tests. In [10] the distance between index finger and thumb is computed for a grasping scenario. Then this distance is combined with actor's current position to compute a X_e during execution of each cycle FM algorithm. There are some limitations of using offline computation for X_e values. These limitations are discussed in Chapter 6.

Representation of the modified FM along with action estimation loop is shown in Figure 5.3.1 with block diagrams. Differences between the FM in [10] and the one in the thesis can be seen by comparing Figure 5.3.1 and Figure 5.2.1.1.

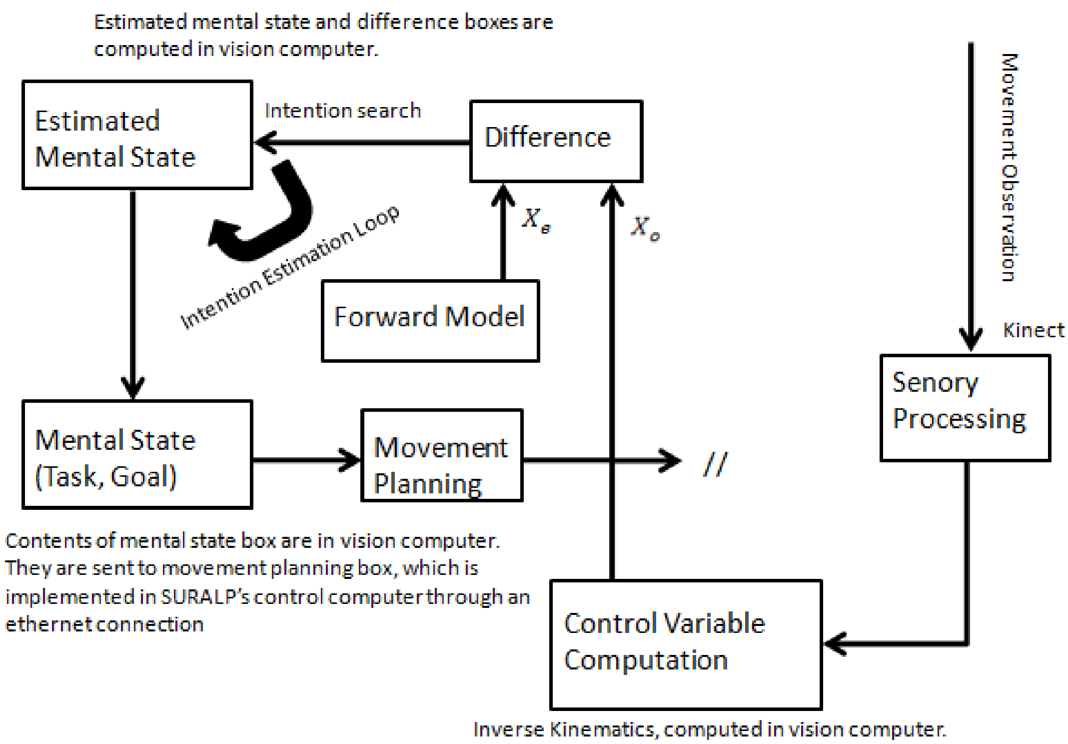

Figure 5.3.1: Block Diagram Representation of the Proposed Intention Estimation Loop. // Denotes that SURALP's Control Computer is not utilized until a decision is made.

There are some important differences between Figure 5.3.1 and Figure 5.2.1.1. In Figure 5.2.1.1 there is a feedback mechanism inside the FM, this model is simplified in Figure 5.3.1. There is no connection between Mental State box and Control Variable Computation box in Figure 5.3.1, though this is not the case in Figure 5.2.1.1. Since inverse kinematics is enough for control variable computation in the simplified model and inverse kinematics can be solved using visual data, Control Variable Computation box do not require inputs from Mental State box.

The differences to the proposed model in [10] significantly simplifies the original model. This simplification can find basis in the fact that animals also posses certain inputs and come to conclusion on intentions of others [39]. It is stated in [40] that animals have simpler models for action estimation than humans.

5.4. Setting

In experiments it is assumed that intentions of an actor are a finite set, so that the FM can make an exhaustive search in the set. In the experimental setting it is accepted that an actor can have three different intentions. These intentions are: (1) approaching robot with an elbow up configuration (2) approaching robot with an elbow down configuration (3) moving hand above the head level. Robot responds to these intentions in order to assist the actor. Robot's respond to intention (1) is to reach with an elbow down configuration, and respond to intention (2) is to reach with an elbow up configuration. If intention (3) is estimated by the FM of robot, it also moves its hand above head level. Figures 5.4.1, 5.4.2 and 5.4.3 show joint angle trajectories of estimated vector of control variables T_e for each intention. Color representations of joint angles trajectories are given in Table 5.2.1.

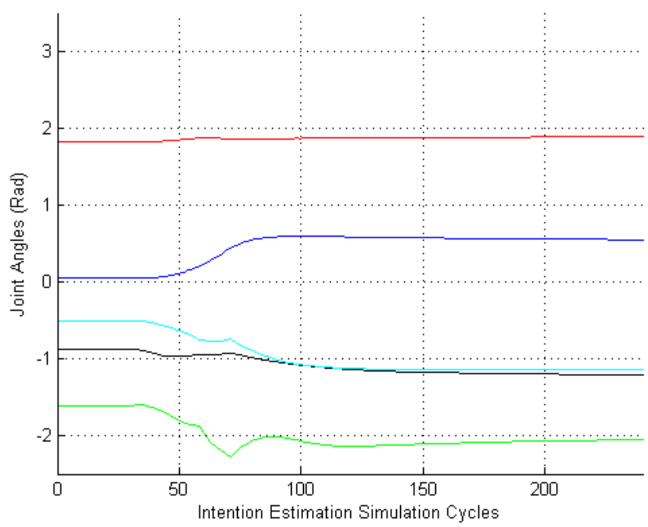

Figure 5.4.1: Joint angle trajectories for intention of reaching with an elbow down configuration, intention (1)

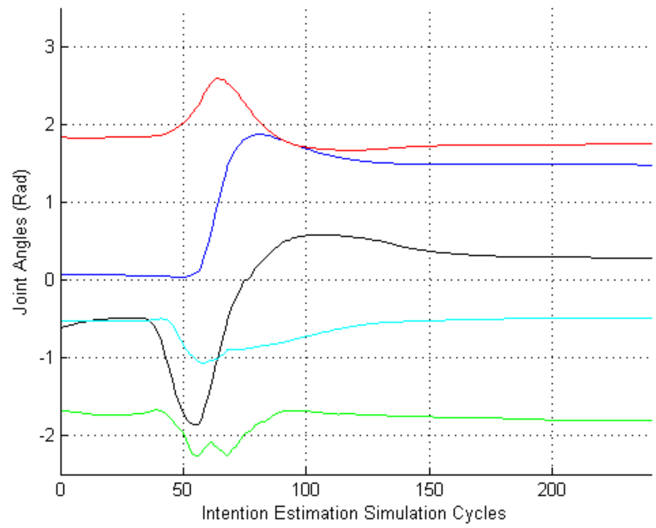

Figure 5.4.2: Joint angle trajectories for intention of reaching with an elbow up configuration, intention (2)

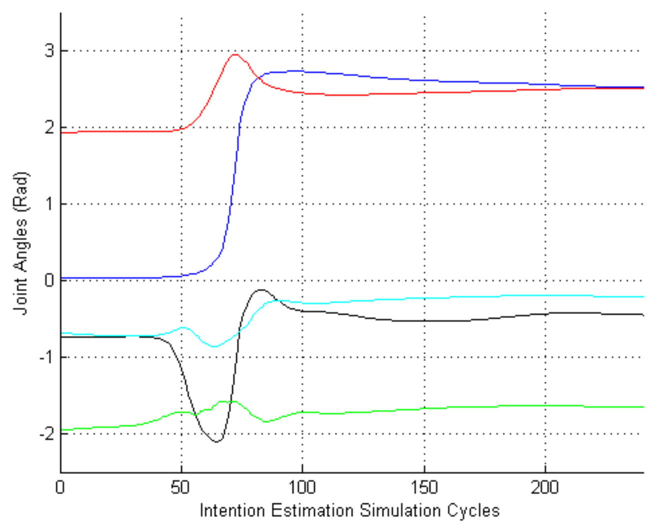

Figure 5.4.3: Joint angle trajectories for intention of moving hand above head level, intention (3)

There are two computers in the technical setting. One of them is the Control Computer of SURALP and the other one is the Action Estimation Computer. Signals which control DC motors connected to each joint are created in the Control Computer. Action Estimation Computer is responsible for processing visual information, storing X_e values and running the intention estimation loop. Once a decision is made it is sent to Control Computer for physical execution. The reason to use two different computers is related to real time computation concerns.

The connection between Control Computer and Action Estimation Computer is established by an Ethernet cable and a communication system composed of a host and client. In the setting host is the Action Estimation Computer and client is the Control Computer. Windows Socket API is used to code host and client applications. An integer number is sent to SURALP control computer at the decision cycle. When this number is received by the control computer, SURALP starts to execute the response to the estimated action. SURALP's responses to different intentions are pre-computed, fixed actions.

A primitive analogy between the technical setting and the biological counterpart can be made at this point. Control computer can be regarded as the primary motor cortex, the part of human brain which executes movements. The Action Estimation Computer be seen as a combination of premotor cortex and parietal cortex of human brain.

One last comment on setting can be made between experimental setup of the proposed technical realization and real time application. In the experimental setup T_o values were also computed offline. Afterwards T_o and T_e values were normalized to a length of 800. In the real time setting Kinect is able to compute an average 20 frames per second and an overall physical execution time of an action is around 3 seconds. Therefore in the real time setting T_o and T_e lengths are around $60\ to\ 80$. T_e plots Figures 5.4.1, 5.4.2 and 5.4.3 are obtained from real time application.

Experimental setup was created to gather fast test results from T_o values. Control Computer was not used, so SURALP was not executing the response to the estimated action. Real time application was implemented after producing satisfactory results in experimental setup. Video capture from real time setting for each intention are shown in Figures 5.4.4, 5.4.5 and 5.4.6.

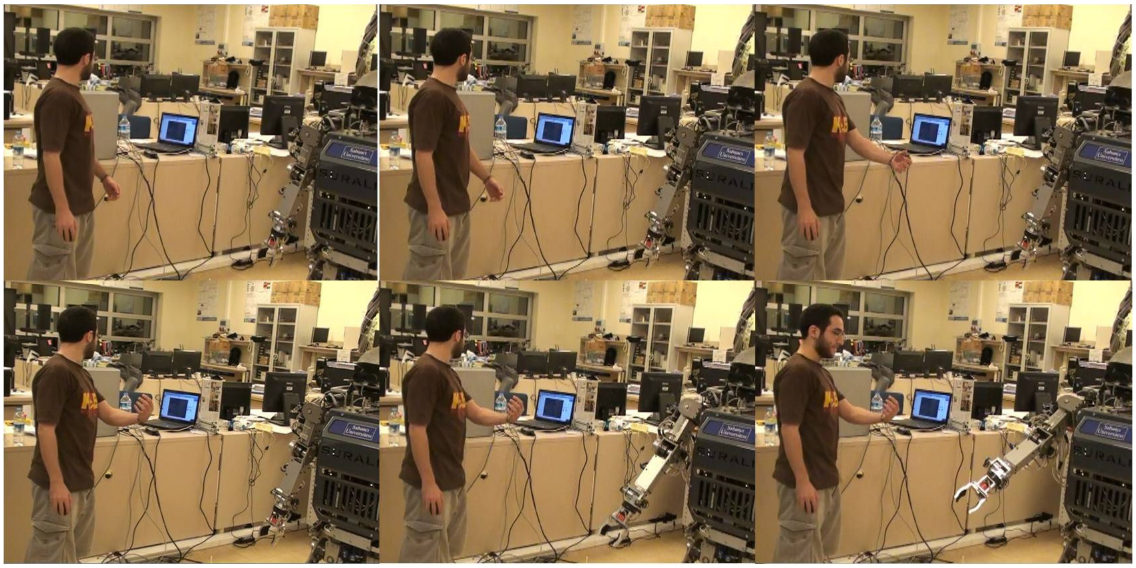

Figure 5.4.4: Elbow down intention in a real time application of estimation intention loop

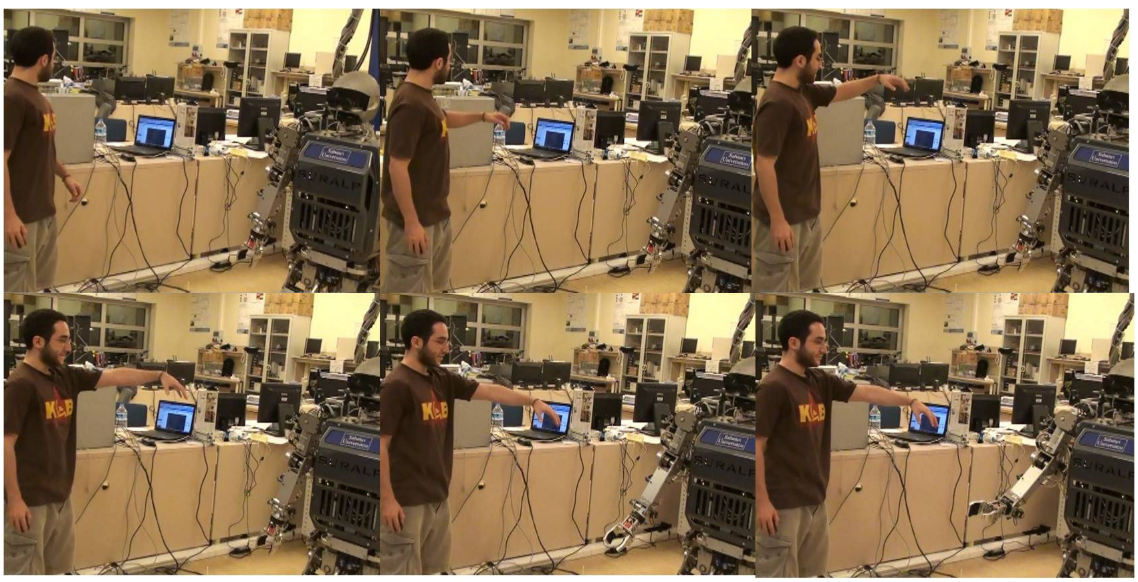

Figure 5.4.5: Elbow up intention in a real time application of estimation intention loop

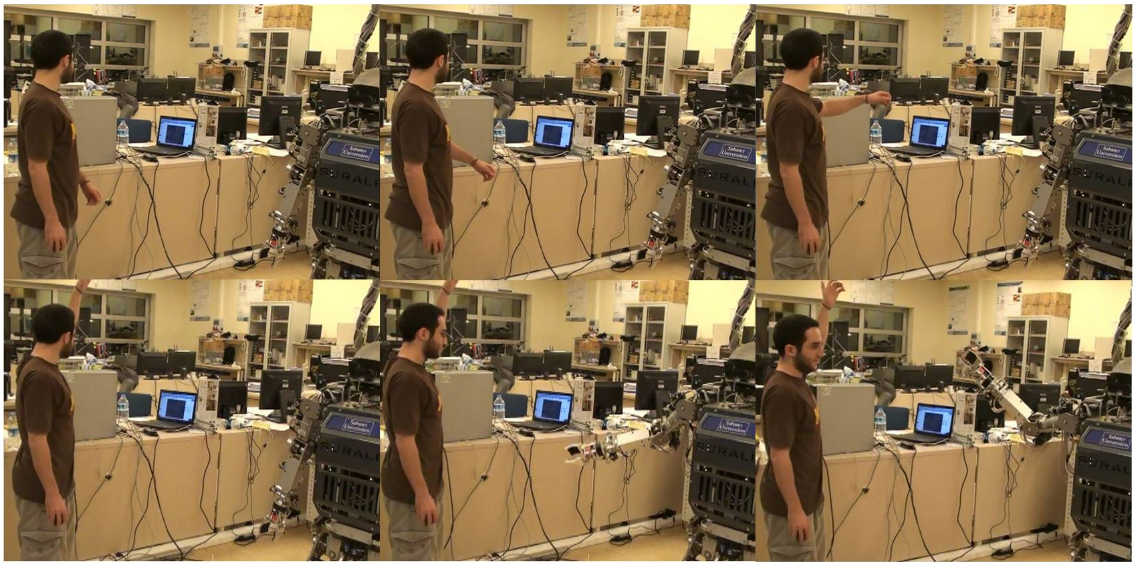

Figure 5.4.6: Intention of moving hand over head level in a real time application of estimation intention loop

5.5. Experiments & Results

Two sets of tests were conducted to test the action estimation loop with the setting proposed in the previous section. There are 24 tests in each one. Each set has 8 movements of each intention. Difference between tests in two sets is the cycle in which a decision is made regarding intentions of actors. In the first set a decision is made in the 250^{th} cycle. In the second set a decision is made in the 100^{th} cycle. It can be observed from Figures 5.5.1, 5.5.2 and 5.5.3 that making a decision at 250^{th} cycle results in robot waiting for nearly the end of execution before executing the estimated intention. On the other hand making a decision in the 100^{th} cycle result in robot to anticipate the actor's intention and execute the proper response to assist the actor.

Table 5.5.1 and Table 5.5.2. show results of these tests. These results indicate that trying to anticipate actor's intention by making a decision at an early cycle increases false decisions. According to these tables making estimation at 250^{th} results in 19 correct estimations, and making estimation at 100^{th} cycle results in 17 correct estimations.

It is proposed by this work to disregard potentially wrong decisions at the cost of some correct decisions. As a result SURALP should be able to make less estimation with less error. From a HMI perspective this can be regarded as a safer working mode. To avoid false results, mean of error between vector of observed and simulated control variables was computed for each test. Mean of error was observed to be higher in false estimations than correct estimations. A threshold value of 0.07 radians was chosen to detect false estimates and classify them as not goal oriented actions.

Table 5.5.1: Results of Action Estimation Tests, Decision Cycle: 250

Intention	**Estimation**	**Mean of Difference**
Elbow Down 1	Correct	0.0394
Elbow Down 2	Correct	0.0125
Elbow Down 3	Correct	0.0148
Elbow Down 4	Correct	0.0486
Elbow Down 5	Correct	0.0260
Elbow Down 6	Correct	0.0239
Elbow Down 7	Correct	0.0092

Elbow Down 8	Correct	0.0263
Elbow Up 1	Correct	0.0266
Elbow Up 2	False (Intention 3 estimated)	0.0327
Elbow Up 3	Correct	0.0266
Elbow Up 4	Correct	0.1215
Elbow Up 5	Correct	0.2374
Elbow Up 6	Correct	0.0520
Elbow Up 7	Correct	0.0596
Elbow Up 8	False (Intention 3 estimated)	0.1016
Moving Hand Above Head 1	Correct	0.0663
Moving Hand Above Head 2	Correct	0.0427
Moving Hand Above Head 3	False (Intention 2 estimated)	0.1222
Moving Hand Above Head 4	False (Intention 2 estimated)	0.0858
Moving Hand Above Head 5	Correct	0.0284
Moving Hand Above Head 6	Correct	0.0571
Moving Hand Above Head 7	False (Intention 2 estimated)	0.0842
Moving Hand Above Head 8	Correct	0.0233

Table 5.5.2: Results of Action Estimation Tests, Decision Cycle: 100

Intention	**Estimation**	**Mean of Difference**
Elbow Down 1	Correct	0.0310
Elbow Down 2	Correct	0.0163
Elbow Down 3	Correct	0.0046
Elbow Down 4	Correct	0.0240
Elbow Down 5	Correct	0.0196
Elbow Down 6	Correct	0.0468
Elbow Down 7	Correct	0.0179
Elbow Down 8	Correct	0.0227
Elbow Up 1	Correct	0.0089
Elbow Up 2	False (Intention 3 estimated)	0.0133
Elbow Up 3	Correct	0.0089

Elbow Up 4	False (Intention 3 estimated)	0.0287
Elbow Up 5	Correct	0.1585
Elbow Up 6	Correct	0.0318
Elbow Up 7	False (Intention 3 estimated)	0.0475
Elbow Up 8	False (Intention 3 estimated)	0.0812
Moving Hand Above Head 1	False (Intention 2 estimated)	0.0215
Moving Hand Above Head 2	Correct	0.0361
Moving Hand Above Head 3	Correct	0.0113
Moving Hand Above Head 4	False (Intention 2 estimated)	0.0789
Moving Hand Above Head 5	Correct	0.0108
Moving Hand Above Head 6	Correct	0.0253
Moving Hand Above Head 7	False (Intention 2 estimated)	0.0928
Moving Hand Above Head 8	Correct	0.0220

An additional set of 8 tests were conducted to test the threshold of 0.07 radians. These tests were not goal oriented actions, i.e. they were not one of three intentions. It can be seen from Table 5.5.3 that all of these tests were decided as not being one of original three intentions.

Table 5.5.3: Results of Tests which were not goal oriented, Decision Cycle: 100

Intention	**Estimation**	**Mean of Difference**
Not a goal oriented action 1	No Estimation	0.3945
Not a goal oriented action 2	No Estimation	0.7523
Not a goal oriented action 3	No Estimation	0.1361
Not a goal oriented action 4	No Estimation	0.1157
Not a goal oriented action 5	No Estimation	0.1431
Not a goal oriented action 6	No Estimation	0.2165
Not a goal oriented action 7	No Estimation	0.3055
Not a goal oriented action 8	No Estimation	0.1522

In both sets no false decisions were observed in intention (1), reaching with an elbow down configuration. Figures 5.5.1, 5.5.2 and 5.5.3 show T_o graphs with respect time for each intention. Intention (2) and intention (3) have similar beginnings. Both actions start their execution by activating shoulder angles. Difference between intentions become more apparent through end of execution when 3 shoulder angles settle on their terminal values. It can be concluded that observer needs to wait until final cycles of mental simulation to make a better distinction between these two intentions. It can also be proposed that intention (2) and (3) are closer to each other in the joint space of SURALP than intention (1).

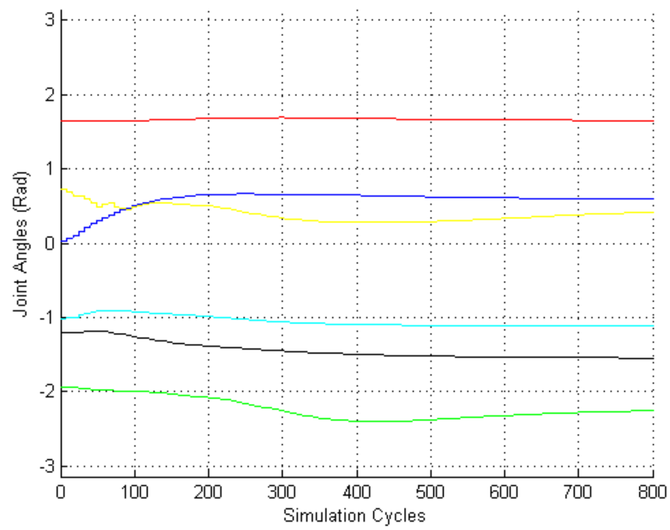

Figure 5.5.1:Joint angle trajectories for Elbow Down 1 test from first set.

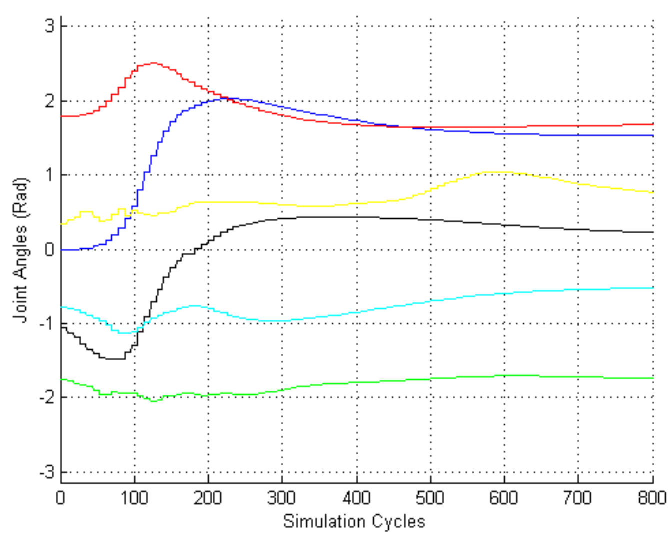

Figure 5.5.2: Joint angle trajectories for Elbow Up 1 test from first set.

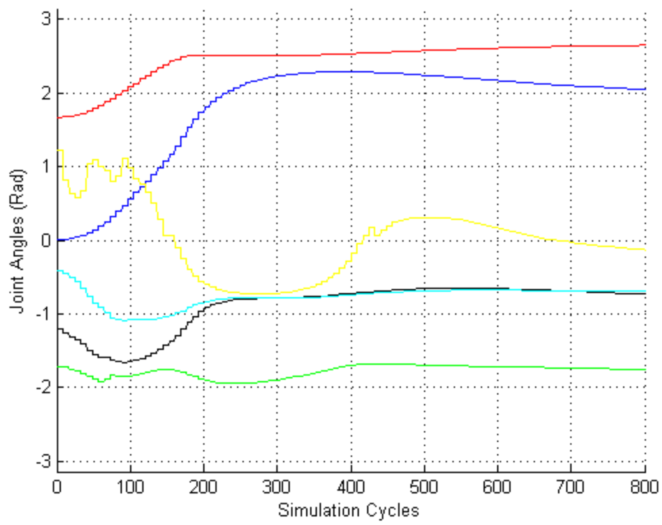

Figure 5.5.3: Joint angle trajectories for Moving Hand Above Head 1 test from first set.

In Figures 5.5.4, 5.5.5 and 5.5.6 probabilities of each intention throughout the simulation are shown. In Figure 5.5.4 and Figure 5.5.5 probability trajectories settle in correct estimation as early as 50^{th} cycle. It can be seen that intention estimation loop is able to distinguish elbow down and elbow up movements in these tests.

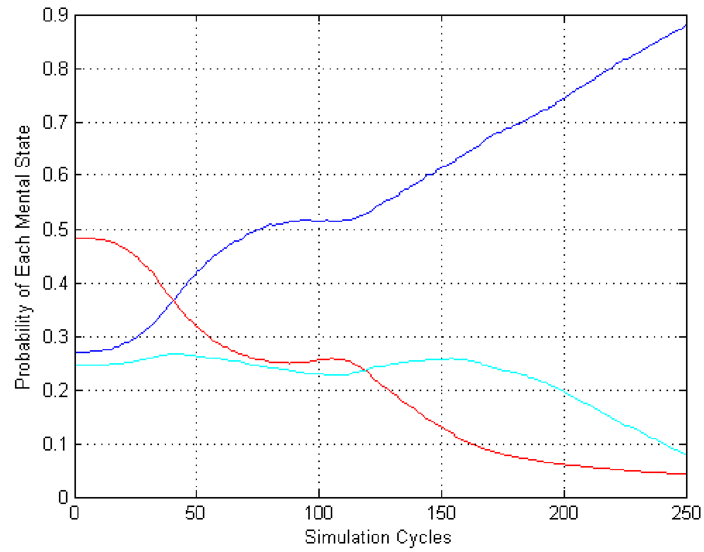

Figure 5.5.4: Probabilities of each intention for Elbow Down 1 test from first set. Blue trajectory is intention (1), red is intention (2), cyan is intention (3).

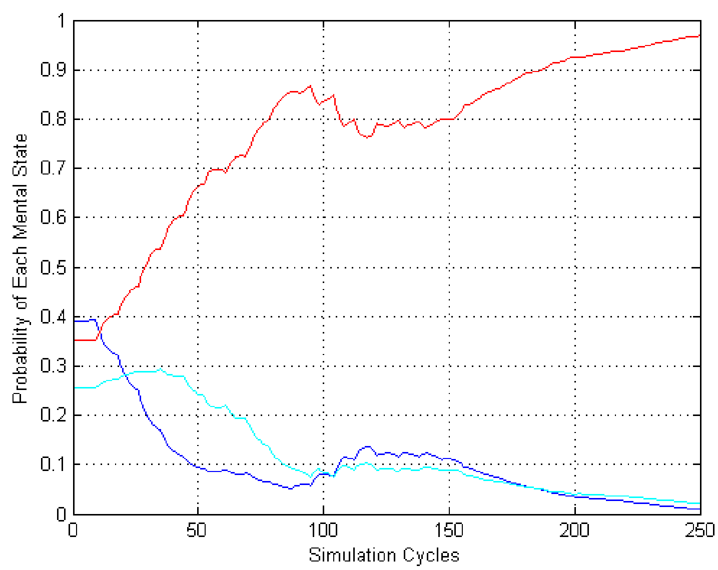

Figure 5.5.5: Probabilities of each intention for Elbow Up 1 test from first set. Blue trajectory is intention (1), red is intention (2), cyan is intention (3).

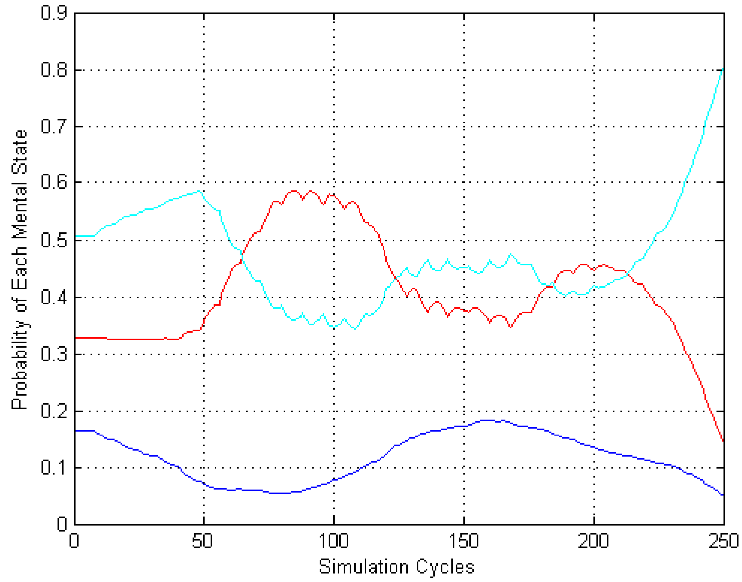

Figure 5.5.6: Probabilities of each intention for Moving Hand Above Head 1 test from first set. Blue trajectory is intention (1), red is intention (2), cyan is intention (3).

Probability plots in Figure 5.5.6 are especially interesting due to the fact that SURALP's estimation keep on changing throughout the simulation. An explanation to this change of belief can be attributed to different frequencies in actions. In other words an action can be performed faster than the stored action in SURALP's FM. In such a scenario SURALP may have unusual estimations regarding the actor's movement. This unusual estimation due to different frequency of actions can explain the oscillation in Figure 5.5.6. Solutions to this limitation are discussed in the next chapter.

Chapter 6

6. Conclusions

A computational FM which was in parallel with findings of contemporary neuroscience literature on action understanding was implemented with success on a robotic platform. There are two key points in this success. The first point is that this work was able to show findings of neuroscience can be applied in robotics, so it can be discussed that robots with better decision making capabilities can be designed by inspiration from human brain models. Robotic researchers can find solutions to problems in the field of HRI from computational theories of human mind. As for the second key point, a robot platform which can anticipate an actor's intentions was realized. Such a robot platform can be used in a setting where humans and robot are required to work together.

Experiment in [48] was explained in Chapter 3. Researchers placed 5 to 12 light sources on human actors in a dark room. Observers were able to estimate human actions from lights sources as low as 5. In this thesis arm actions of humans were estimated from 5 joint angles. It can be concluded that findings of thesis are in parallel with [48].

6.1. Limitations of the FM in the Thesis

Implementation of the computational FM algorithm was successful, but it has limitations. One of the limitations is related to computation of X_e, vector of simulated control variables. Another limitation is the speed of movements. As explained in the last chapter current FM have no method for anticipating an action unless observed and simulated actions have same speeds. The last limitation is based on neurological motivation of the FM.

Computing X_e, vectors of control variables, is a key issue. In this work X_e is computed by solving inverse kinematics problem offline, meaning that values of X_e are same for all action estimation tests. It can be put forward that a dynamic X_e generator, which creates different T_e matrices for each run of action estimation algorithm can lead to higher success rates. In such a scenario FM should not contain the time trajectory of an action but some clues about the action. For reaching tasks with elbow down and elbow up configurations these clues can be the height difference between hand and elbow of the actor, and the distance between end effector and actor's hand. Then FM can create a $X_e$for the current cycle by combining clue with current positions of the actor. It should be pointed out that a dynamic computation for X_e is proposed in the simulations of [10], but details of computing $X_e$from current observer position and clues regarding the intentions are omitted.

One of the problems of using pre computed X_e is that two same sequences of actions might have different lengths. For example, when an elbow down action is executed slowly it produces a longer T_o vector. The FM in this thesis would not be able to detect such actions due to pre computation of T_e. A dynamic FM as explained in the last paragraph can also solve this issue. Moreover a solution to this problem can also be implemented by using Dynamic Time Warping algorithm [66], a method for measuring similarities between two vectors of different lengths.

Last limitation is related to mirror neurons. The computational FM proposed in [10] extensively gives references to mirror neurons, but mirror neurons itself is a very new topic in neuroscience and there are critiques [64, 65, 67]. These critiques argue on questions regarding origins of mirror neurons and applications of mirror neurons to action estimating models. [68] argues against simulation capabilities of mirror neurons in a philosophical context. If these critiques are proved to be right in the future, computational model for estimating actions would lose its neurological basis. Nevertheless model is capable of estimating, even anticipating, an actor's intentions from a discrete set within reasonable time without its neurological basis.

6.2. Possible Improvement on the FM

It was reported in Chapter 5 that SURALP was using pre-computed trajectories for responding to actors. Dynamic trajectories for SURALP's response can be computed using the distance information between SURALP's end effector and actor's hand. It was also reported that the computational FM has dual function. During action observation it is used to make estimations and during action execution it is used to reduce sensory delays by anticipating visual results of actions. It is not verified in [008] whether such a FM can reduce delays. To improve the computational FM implemented in this thesis, same computational

FM in the vision computer can be used to simulate the action of the actor until the decision cycle. After the decision cycle same FM can be triggered to anticipate visual consequences of reaching the actor's hand and uses the data to reduce delays. These trajectories can be compared with the dynamic responses obtained without FM to observe if delays are actually reduced as proposed in [008].

6.3. Future Work

In the thesis a lot of attention is given to mirror based learning mechanisms of human mind. On the other hand proposed solution to action estimation in the thesis does not use a mirror based learning approach. Most of the parameters are tuned offline: A decision cycle, creation of estimated actions, differentiation threshold for action detection. This learning mechanism is especially required during creation of different sets of actions. A clustering analysis can be made to organize joint angle trajectories into different sets. Such a clustering analysis can be made without supervision from human operations. After clustering of data is complete proposed FM and intention estimation loop can be run.

Not overlapping joint angle trajectories are a serious problem in decision making process. This problem can be solved in three ways. Same intentions with different lengths can be added to the estimated intentions set in robot's mind. Therefore intention estimation loop can detect slower and faster versions of the same intention. Another method way of solving the problem can be implementation of an algorithm called Dynamic Time Warping. This algorithm dynamically normalizes two sequences and computes the similarity between them. A dynamic FM can be implemented. This model can produce estimated control vectors with respect to current observed visual data. Such a dynamic FM could also be a solution this problem.

Detection of certain action sets depend more on certain joints. For example sets in this thesis depend on shoulder and elbow angle joint trajectories. A reaching action with a certain hand orientation on the other hand is expected to depend also on joints which directly rotate the end effector. If more important joints in an action set are determined from the sets in robot's mind, action joint trajectories can be stored in with fewer joint angles. A way to detect parameters which have less effect on decision making is Principle Component Analysis (PCA). It can be used after estimated joint angle trajectories in robot's mind are computed and organized with a clustering algorithm.

After improving the intention estimation loop with proposed methods, its success with larger intention sets needs to be observed. An action intention set with more than 3 elements can be organized and larger test data can be used to determine the success of the intention estimation loop.

REFERENCES

[1] V. Gallesse, L. Fadiga, L. Fogassi, G. Rizzolatti. "Action Recognition in the Premotor Cortex", *Brain a Journal of Neurology,* Vol. 119, Issue 2, pp: 593-609, October 1995

[2] F. P. de Lange, M. Spronk, R. M. Willems, I. Toni, H. Bekkering. "Complementary Systems for Understanding Action Intentions," *Current Biology,* Vol. 18, Issue 6, pp: 454-57, February 2008

[3] M. Thioux, V. Gazzola, C. Keysers, "Action Understanding: How, What and Why," *Current Biology*, Vol. 18, Issue 10, pp: 431-34, May 2008

[4] G. Pellegrino, L. Fadiga, L. Fogassi, V. Gallesse and G. Rizzolatti: "Understanding Motor Events: A Neurophysiological study", *Experimental Brain Research,* Vol. 91, Issue 1, pp. 176-180, September 1992.

[5] G. Rizzolatti, L. Craighero. "The Mirror Neuron System", *Annual Review of Neuroscience,* Vol. 27, pp. 169-92, 2004

[6] L. M. Oberman, V. S. Ramachandran. "Reflections on the Mirror Neuron System Their Evolutionary Functions Beyond Motor Representation," in *Mirror Neuron Systems,* J. A. Pineda, Humana Press, 2009, pp: 39-59

[7] J. Järveläinen, M. Schürmann, R. Hari. "Activation of the human primary motor cortex during observation of tool use", *NeuroImage*, Vol. 23, Issue 1, pp: 187-192, September 2004

[8] E. S. Cross, A. F. de C. Hamilton, S. T. Grafton, "Building a motor simulation de novo: Observation of dance by dancers", *NeuroImage*, Vol. 31, Issue 3, pp: 1257-1267, July 2006

[9] K. J. Wheaton, J. C. Thompson, A. Syngeniotis, D. F. Abbott, A. Puce, "Viewing the motion of human body parts activates different regions of premotor, temporal, and parietal cortex", *NeuroImage*, Vol 22, Issue 1, pp: 277-288, May 2004

[10] E. Oztop, D. Wolpert, M. Kawato, "Mental state inference using visual control parameters", *Cognitive Brain Research*, Vol. 22, Issue 2, pp: 129-51, February 2005

[11] C. L. Baker, J. Tenenbaum, R. Saxe. "Bayesian models of human action understanding," *in Neural Information Processing Systems*, 2006

[12] C.L. Baker, R. Saxe, J.B. Tenenbaum. "Action Understanding as inverse planning," *Cognition,* Vol. 113, Issue 3, pp: 329-49, December 2009

[13] P. Smagt, M. Grebenstein, H. Urbanek, N. Fligge, M. Strohmayr, G. Stillfried, J Parrish, A. Gustus. "Robotics of human Movements," *Journal of Physiology,* pp: 119-32, 2009

[14] P. Davidsson, E. Astor, B. Ekdahl. “A Freamework for Autonomous Agents Based on the Concept of Anticipatory Systems,” *In Cybernetics and Systems,* 1994

[15] T.C. Lueth; T. Laengle; G. Herzog; E. Stopp; U. Rembold, "KANTRA-human-machine interaction for intelligent robots using natural language," *Robot and Human Communication, 1994. RO-MAN '94 Nagoya, Proceedings., 3rd IEEE International Workshop on* , pp. 106-111, July 1994

[16] J.S. Albus. "Outline for a theory of intelligence," *Systems, Man and Cybernetics, IEEE Transactions on ,* Vol. 21, no.3, pp: 473-509, May 1991

[17] H. Kazerooni "Extender: a case study for human-robot interaction via transfer of power and information signals," *Robot and Human Communication, Proceedings., 2nd IEEE International Workshop on* , pp: 10-20, November 1993

[18] W.Y. Albert, Y.K. Henry. “Intelligent Robot-assisted Humanitarian Search and Rescue System”, *International Journal of Advanced Robotic Systems,* pp: 121-28, 2009

[19] H. Kozima, C. Nakagawa, Y. Yasuda, “Children–robot interaction: a pilot study in autism therapy,” *Progress in Brain Research*, Elsevier, Vol. 164, pp: 385-400, 2007

[20] K. Dautenhahn, I Werry. “Issues of Robot-Human Interaction Dynamics in the Rehabilitation of Children with Autism,” *In Proc. of FROM ANIMALS TO ANIMATS, The Sixth International Conference on the Simulation of Adaptive Behavior,* 2000

[21] . Robins, K. Dautenhahn, J. Dubowski. "Investigating autistic children's attitudes towards strangers with the theatrical robot - a new experimental paradigm in human-robot interaction studies," *Robot and Human Interactive Communication, 2004. ROMAN 2004. 13th IEEE International Workshop* on, pp.557-62, September 2004

[22] D. Seifer, M. Mataric. “Robot-assisted therapy for children with Autism Spectrum Disorders,” *Proceedings of the 7th international conference on Interaction design and children*, pp: 49-52, 2008

[23] C. Breazeal. “Emotion and sociable humanoid robots,” *International Journal of Human Computer Studies,* Vol. 59, Issue 1-2, pp: 119-55, July 2003

[24] M. Yılmaz, U. Seven, K. Erbatur. “Biped robot walking control on inclined planes with fuzzy parameter adaptation,” *In: IFAC Workshop Adaptation and Learning in Control and Signal Processing*, 2010

[25] M. Fujita, K. Kageyama, “An open architecture for robot entertainment,” *Proceedings of the first international conference on Autonomous agents,* pp: 435-42, 1997

[26] K. Hornfeck, Y. Zhang, K. Lee. “Philos A Sociable Robot for Human Robot Interactions and Wireless Health Monitoring,” *Proceedings of the 27th Annual ACM Symposium on Applied Computing,* pp: 293-94, March 2012

[27] M. Raibert. “BigDog, the Rough-Terrain quadruped robot,”*Proceedings of the 17th IFAC World Congress,* Vol. 17, 2008

[28] J.E. Pratt, B.T. Krupp, C.J. Morse, S.H. Collins, "The RoboKnee: an exoskeleton for enhancing strength and endurance during walking," *Robotics and Automation, 2004. Proceedings. ICRA '04. 2004 IEEE International Conference on* , Vol. 3, pp: 2430-35, May 2004

[29] S. Stan, M. Manic, V. Maties, R. Balan, "Evolutionary approach to optimal design of 3 DOF translation exoskeleton and medical parallel robots," *Human System Interactions, 2008 Conference on* , pp: 720-25, May 2008

[30] K. E. Gordon, D. P. Ferris, “Learning to walk with a robotic ankle exoskeleton,” *Journal of Biomechanics*, Vol. 40, Issue 12, pp: 2636-44, 2007

[31] C. Schmitt, P. Metrailler. “The motion maker a rehabilitation system combining an orthosis with closed-loop electrical muscle stimulation”, *8th Vienna International Workshop on Functional Electrical Stimulation*, pp: 117-20, 2004

[32] M. Bouri, B. Le Gall, R. Clavel. “A new concept of parallel robot for rehabilitation and fitness the Lambda,” *EEE International Conference on Robotics and Biomimetics,* December 2009

[33] H. S. Koppula, A. Saxena. “Anticipating Human Activities using Object Affordances for Reactive Robotic Response”, *Robotics: Science and Systems,* 2013

[34] Z. Wang, M. Deisenroth, H. B. Amor, D. Vogt, B. Scholkopf, J. Peters. “Probabilitistic Modelling of human Movements for Intention Inference,” *Proceedings of Robotics: Science and Systems,* July 2012

[35] C. Frith, U. Frith. “Interacting Minds - A Biological Basis,” *Science Magazine*, Vol. 286, No 5445, pp: 1692-95, 1999

[36] V. Gallese, A. Goldma. “Mirror neurons and the simulation theory of mind-reading”, *Trends in Cognitive Sciences*, Vol. 2, Issue 12, pp: 493-501, December 1998

[37] S.J. Blakemore, J. Decety. “From the Perception of Action to the Understanding of Intention”, *Nature Reviews Neuroscience*,Vol. 2, Issue 8, 561-67, July 2001

[38] C. Lamm, M. H. Fischer, J. Decety, “Predicting the actions of others taps into one's own somatosensory representations—A functional MRI study”, *Neuropsychologia*, Vol 45, Issue 11, pp: 2480-2491, 2007

[39] D. Premack, G. Woodruff. “Does the Chimpanzee Have a Theory of Mind”, *Behavioral and Brain Sciences*, Vol. 1, Issue 4, pp: 515-26, December 1978

[40] J. Call, M. Tomasello. “A Nonverbal False Belief Task The Performance of Children and Great Apes”, *Child Development,* Vol. 90, Issue 2, pp: 381-95, March 1999,

[41] S. Pinker. “How the Mind Works?,” *Annals of the New York Academy of Sciences,* Vol. 882, pp: 119-27, 1999

[42] E. Kohler, C. Keysers, M.A. Umilta, L. Fogassi, V. Gallesse, G. Rizzolatti. “Hearin Sounds, Understanding Actions Action Representation in Mirror Neurons,” *Science Magazine,* Vol. 297, No 5582, pp: 846-48, August 2002

[43] M. Iacoboni, I. Molnar-Szakacs, V. Gallese, G. Buccino, J.C: Mazziotta. “Grasping the Intentions of Others with One's Own Mirror Neuron System,” *PLoS Biol 3(3): e79*, 2005

[44] M. Dapretto, M. Davies, J. Pfeifer, A. Scott, M. Sigman, S. Bookheimer, M. Iacoboni. “Understanding emotions in others mirror neuron dysfunction in children with autism spectrum disorders,” *Nature Neuroscience,* 9, pp: 28-30, December 2005

[45] G. Rizzolatti, L. Fogassi, V. Gallesse. “Neurophysiological mechanisms underlying the understanding and imitation of action,” *Nature Reviews Neuroscience,* Vol. 2, 660-70, September 2001

[46] E. Oztop, M. Kawato, M. Arbib, “Mirror neurons and imitation: A computationally guided review,” *Neural Networks*, Vol. 19, Issue 3, pp: 254-71, April 2006

[47] J. Zhong, C. Weber, S. Wermter “Robot Trajectory Prediction and Recognition based on a Computational Mirror Neurons Model,” *Artificial Neural Networks and Machine Learning – ICANN,* 2011

[48] G. Johansson. “Visual Perception of Biological Motion and a Model for its Analysis”, *Perception & Psychophysics*, Vol 14, Issue 2, pp: 201-11, 1973

[49] D. Bernstein, K. Crowley, I. Nourbakhsh. “Working with a robot exploring relationship potential in human-robot systems”*, Interaction Studies*, Vol.8, Issue 3, pp: 456-82, 2007

[50] A. Edsinger, C.C. Kemp, "Human-Robot Interaction for Cooperative Manipulation: Handing Objects to One Another," *Robot and Human interactive Communication*, pp: 1167-72, August 2007

[51] G. Hoffman, C. Breazeal. “Collaboration in Human-Robot Teams,” *In Proc. of the AIAA 1st Intelligent Systems Technical Conference,* 2004

[52] D. Dementhon, L. Davis. "Model-Based Object Pose in 25 Lines of Code," *International Journal on Computer Vision,* Vol. 15, Issue 12, pp: 123-41, 1995

[53] B. Rosehahn, C. Perwass, G. Sommer. "Pose Estimation of 3D Free-Form Contours," *International Journal on Computer Vision,* Vol. 62, Issue 3, pp: 267-89, 2005

[54] P. Kohli, J. Rihan, M. Bray, P. Torr. "Simultaneous Segmentation and Pose Estimation of Humans Using Dynamic Graph Cuts," *International Journal on Computer Vision,* Vol. 79, Issue 3, pp: 285-98, 2008

[55] P. Kohli, J. Shotton. "Key Developments in Human Pose Estimation for Kinect," in *Consumer Depth Cameras for Computer Vision*, A. Fossati, J. Gall, H. Grabner, X. Ren, K. Konolige, Springer, 2013, pp: 63-70

[56] M. Yilmaz, U. Seven, K.C. Fidan, T. Akbas, K. Erbatur. "Circular arc-shaped walking trajectory generation for bipedal humanoid robots," *Advanced Motion Control (AMC), 2012 12th IEEE International Workshop on* , pp: 1-8, March 2012

[57] M. Yılmaz, U. Seven, K. Erbatur. "Biped robot walking control on inclined planes with fuzzy parameter adaptation," *In: IFAC Workshop Adaptation and Learning in Control and Signal Processing*, 2010

[58] S. H. Johnson, P. M. Corballis, M. S. Gazzaniga, "Within grasp but out of reach: evidence for a double dissociation between imagined hand and arm movements in the left cerebral hemisphere," *Neuropsychologia*, Vol. 39, Issue 1, pp 36-50, January 2001

[59] J. Jastorff, C. Begliomini, M. Destro, G. Rizzolatti, GA. Orban. "Coding Observed Motor Acts Different Organizational Principles in the Parietal and Premotor Cortex of Humans," *Journal of Neurophysiology*, Vol. 104, Issue 1, pp: 128-40, Jul 2010

[60] Y. Burnord, P. Baraduc, A. Battaglia, E. Guigon, E Koechlin, S. Ferraina, F. Lacquaniti, R. Caminiti, "Parieto-frontal coding of reaching an integrated framework," *Experimental Brain Research,* Vol. 129, Issue 3, pp: 325-46, December 1999

[61] R. A. Andersen. "Multimodal Integration for the Representation of Space in the Posterior Parietal Cortex," *Philosophical Transactions of the Royal Society*, Vol. 352, no. 1360, pp: 1421-28, October 1997

[62] P. Fattori, M. Gamberini, DF Kutz, C. Galletti. "Arm-reaching neurons in the parietal area V6A of the macaque monkey," *European Journal of Neuroscience,* Vol. 13, Issue 12, pp: 2309-11, Jun 2001

[63] D. Purves, D. Fitzpatrick, W. C: Hall, A. S. Lamantia, L. E. White. *Neuroscience,* Sinauer Associates, 2012.

[64] V. Kosonogov. "Why the Mirror Neurons Cannot Support Action Understanding", *Neurophysiology*, Vol. 44, No. 6, pp: 499-502, December 2012

[65] G. Hickok. "Eight problems for the mirror neuron theory of action understanding in monkeys and humans", *Journal of Cognitive Neuroscience*, Vol. 21, Issue 7, pp: 1229-43, July 2009

[66] S. Salvadore, P. Chan "FastDTW: Toward accurate dynamic time warping in linear time and space" *In 3rd Workshop on Mining Temporal and Sequential Data,* 2004

[67] C. Heyes, "Where do mirror neurons come from?", *Neuroscience & Biobehavioral Reviews*, Vol. 34, Issue 4, pp: 575-83, March 2010

[68] B.V. Hole. "Mirror Neurons, Mental Simulation and Mind-Reading," *24th International Workshop on Qualitative Mind Reading,* 2010