SOCIALLY ASSISTIVE CHILD-ROBOT INTERACTION IN PHYSICAL EXERCISE COACHING

by

Arzu Güneysu

B.S., Computer Engineering, Boğaziçi University, 2013

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Computer Engineering
Boğaziçi University
2016

ACKNOWLEDGEMENTS

I would like to thank my parents, Nuh and Elmas Güneysu for their invaluable support and endless love.

I would like to express my deepest gratitude to Ayberk Özgür not only for being my love and best friend but also for his encouragement and support.

I would like to extend my appreciation to my thesis advisor Prof. Bert Arnrich for helping me to set realistic research goals on this interdisciplinary subject and guiding me on my way to accomplish my academic goals.

I would like to thank Prof. Levent Akın for giving me to chance to working on robots and his support to do interdisciplinary social robotics studies during my education in Boğaziçi University.

I thank Prof. Albert Ali Salah for his guidance and Recep Doğa Siyli for his great contribution and motivation through the study of offline evaluation system and experiments with children.

I thank Prof. Özlem Durmaz İncel and Prof. Arzucan Özgür for being members of my thesis committee.

I would like to thank exercise coach Ozan Atlı for his great effort to contribute to my study with his professional background on exercise coaching and physiology. I am grateful to Eren Tutkun, Emre Tutkun, Asya Baştürk, Mertcan Baştürk, Deniz Aytuğ for their invaluable help and participation in our several preliminary data collection studies. Their great interest and excitement during the social robot studies always made me inspired and pushed me to do a better system design in each step. Also I would like to thank all child participants of our data collection study for their great interest in my interaction system and for their invaluable feedback.

I thank Yasemin Usta, Can Tunca, Okan Aşık, Yiğit Yıldırım and Bahar Hilal Yüksel for their great help in different phases of my study. I thank Dilbade Rehabilitation Center and Armada Schools for their collaboration and giving permission to observe their exercise therapies and lessons.

The research procedure and the experiments were approved by the ethical committee of Bogazici University. Consents of subjects and consents of their parents were taken for the experiment and for the usage of the visual materials in the thesis and academic papers.

This work was partially funded by the Brain Circulation Scheme Project "Pervasive Healthcare: Towards Computational Networked Life Science" (TUBITAK Co-Circ 2236, Grant agreement number: 112C005) supported by TUBITAK and EC FP7 Marie Curie Action COFUND.

ABSTRACT

SOCIALLY ASSISTIVE CHILD-ROBOT INTERACTION IN PHYSICAL EXERCISE COACHING

The main contribution of this thesis is the design and implementation of an autonomous socially assistive human robot interaction system to engage children in performing several physical exercise motions by providing real-time feedback, guidance, and encouragement.

Preliminary offline tests were done within a rehabilitation therapy context but since the potential of the system to encourage such exercise activities is beyond the physical therapy fields, we designed an online system to behave as an exercise coach that can motivate children to start and continue improving their health and become physically active.

The online system was tested in several preliminary experiments with an exercise coach and children during the design process. Furthermore, an experimental study was conducted with 19 healthy children aged between 4 and 12 to test the feasibility and the effectiveness of our online Socially Assistive Robotic (SAR) exercise system across a variety of performance and evaluation measures.

The results of the study validate the effectiveness of the system in motivating and helping children to complete physical exercises. The children engaged in physical exercise throughout the interaction sessions with a robot coach and rated the interaction highly in terms of enjoyableness, and rated the robot exercise coach highly in terms of social attraction, social presence, and companionship via a questionnaire answered after each session.

ÖZET

FİZİKSEL EGZERSİZ KOÇLUĞU İÇİN SOSYAL DESTEKLİ ÇOCUK ROBOT ETKİLEŞİMİ

Bu tezin ana katkısı çeşitli fiziksel egzersiz hareketlerini gerçekleştirirken çocukları, gerçek zamanlı geri bildirim, rehberlik ve teşvik sağlayarak motive eden bir sosyal destekli insan robot etkileşimi sisteminin tasarımı ve kodlanmasıdır.

Ön testler rehabilitasyon tedavisi kapsamında yapılmış olmasına rağmen, bu tarz egzersiz faaliyetlerini teşvik etme sisteminin potansiyelinin fizik tedavi alanının ötesinde olmasından ötürü, gerçek zamanlı çalışan sistem, çocukları sağlıklarını iyileştirme için ve fiziksel olarak aktif olmaları için motive eden bir egzersiz koçu gibi davranacak şekilde tasarlanmıştır.

Gerçek zamanlı sistem tasarım aşamasında egzersiz koçu ve çocuklar eşliğinde yapılan bir çok ön çalışma ile test edilmiştir. Ayrıca, gerçek zamanlı sosyal destekli robotik egzersiz sisteminin etkinliğini ve fizibilitesini çeşitli performans ve değerlendirme ölçüleri aracılığla test etmek için 4 ve 12 yaş arası 19 sağlıklı çocuk ile deneysel bir çalışma yürütülmüştür.

Çalışmanın sonuçları sistemin, çocukları fiziksel egzersizleri tamamlamaları için motive etme ve çocukların fiziksel egzersizleri tamamlamalarına yardım etme konularındaki etkinliğini doğrulamıştır. Çocuklar robot antrenör eşliğindeki etkileşim oturumları boyunca fiziksel egzersiz yapmış olup, etkileşim sonrası sorulan anket aracılığıyla, etkileşimi; eğlencelilik açısından, egzersiz robotunu; sosyal sempati, arkadaşlık ve toplumsal varlık açılarından olumlu ve/veya yüksek notlarla puanlamışlardır.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

q_i	X component of the vector part of quaternion
q_j	Y component of the vector part of quaternion
q_k	Z component of the vector part of quaternion
q_r	The scalar part of quaternion
R	Quaternion-derived rotation matrix

LIST OF ACRONYMS/ABBREVIATIONS

3D	Three Dimensional
ASD	Autism Spectrum Disorder
DOF	Degrees of Freedom
DTW	Dynamic Time Warping
FP	False Positive
FN	False Negative
IMU	Inertial Measurement Unit
ICC	Intraclass Correlation Coefficient
SAR	Socially Assistive Robotics
TP	True Positive
TN	True Negative

1. INTRODUCTION

In Socially Assistive Robotics (SAR), the robots provide assistance to the user through an effective interaction where the purpose is to achieve measurable progress in rehabilitation, learning, etc. [1]. Constantly increasing interest in humanoid robotics is motivating novel SAR solutions with humanoids for personal assistance, elderly care, child education and social well-being [2].

Social humanoid robots that can interact verbally and can perform exercise along with a human hold promise in helping to increase physical activity among children and adults within the context of physical rehabilition and exercise coaching. Among these promising areas where SAR can provide solutions, motivating children is an especially important problem since the repetitive nature of the physical rehabilitation and exercise activities decreases continuous attention for the current activity which may lead to decrease in the performance of the child. Since playing a game is a concrete way for children to communicate and express themselves more naturally [3], children's motivation to participate in such exercise activities can be improved by game-like child-robot interaction. Belpaeme et al. showed that physical robots can effectively grasp the attention of children and that adaptation to user characteristics can be a useful tool in supporting long term interaction [4]. In an interaction scenario, the humanoid robot can perform selected exercises or rehabilitation motions and ask the child to imitate itself. Recognizing the actions of the child and serving appropriate feedback is one of the most important features that a robotic system should have in order to achieve an interactive engagement that catches the child's attention and motivates the child to imitate the robot during such an exercise session.

The main contribution of this work is the design and implementation of an autonomous socially assistive human-robot interaction system to boost the engagement of children to perform several physical exercise motions by providing real-time feedback, guidance, and encouragement. The proposed online system is composed of a socially assistive humanoid robot that monitors the performance of the child via an external monitoring system and data processing unit in an exercise scenario, with the purpose of motivating the child to perform the shown motion and improving or correcting the performance via verbal or physical feedback.

In our preliminary studies we designed and tested an offline evaluation system for grading the upper arm gestures of a child during a robotic physiotherapy coaching session with the humanoid robot Nao (see Figure 3.1). We used a Kinect camera to extract the skeleton data of the child to do motion evaluation. Compared to the other studies in literature, instead of recognizing the type of action performed by the child, our proposed system attempts to automatize the evaluation process by mimicking the decision criterias of physiotherapists.

After initial experiments with healthy children where they mimic the actions of a humanoid robot which consist of three exercise motions, our evaluation measure was shown to be reliable according to Intraclass Correlation Coefficient (ICC) when compared with 5 physiotherapists' evaluations as ground truth data. The degree of consistency among our ratings and the physiotherapist ratings is between 76% and 96% for different motions. The results show that physiotherapy motions in the literature can be misunderstood by the children and proper feedback should be given to the children during the exercise.

Then, we designed goal based functional activities in order to help children to understand and perform the motions easily. We observed many theraphy exercises in therapy centers to test our proposed system with Kinect and the robot. However, a limitation was detected during our observations during this study: The therapist stays in front of the children and gives them physical objects like toys during a functional activity and this creates a barrier to Kinect vision. Unfortunately, the Kinect version of the system was not tested in the theraphy centers. Therefore, we decided to use wireless Inertial Measurement Units (IMUs) as an alternative way to monitor motions in real-time. Data from children and physiotherapists were recorded and each motion is analyzed. In order to observe how children react to the robot and to the wearable IMU modules, we tested the system in a rehabilitation center for hemiparetic and diparetic children.

Since the potential of the system to encourage and train such exercise activities is beyond the physical therapy fields, we then designed an online system to behave as an exercise coach that can motivate children to start and continue improving their health and become physically active. The online system is tested with an exercise coach and children during the design process.

An experimental study was conducted with 19 healthy children aged between 4 and 12 to test the feasibility and effectiveness of our online SAR exercise system across a variety of user performance and evaluation measures. The results of the study validate the effectiveness of the system in motivating and helping children to complete physical exercises. The children engaged in physical exercise throughout the interaction sessions and rated the interaction highly in terms of enjoyableness, and rated the robot exercise coach highly in terms of social attraction, social presence, and companionship via a questionnaire.

2. RELATED WORK

Robots have been used to serve various purposes, tasks and applications in the medical/healthcare and social care domain [5]. In the medical healthcare domain, there are many SAR researchers who propose solutions to rehabilitation. Often, the aim is to enhance the ranges of motion and the use of an impaired limb in daily activity. The feasibility of utilizing a humanoid robot in rehabilitation is shown by Choe *et. al.* [6]. Fridin *et. al.* proposed a human-robot interaction scheme with Nao robot for children with Cerebral Palsy that motivates them to participate in the therapy activity [7].

Brooks *et. al.* proposed solutions to quantify upper arm rehabilitation metrics (the range of motion and the peak angular velocity) for children during child-robot play-integrated therapy. In the initial experiments they evaluated their methodology with two easy upper arm exercise motions by directly asking healthy subjects to perform the actions and recording the motions via a webcam [8].

Strohrmann *et. al.* performed a longitudinal study with children that are in need of physiotherapy. The children performed 10 predefined motor tasks such as turning around cards, picking up small objects and climbing stairs. The movements are monitored using wearable sensors and video recordings were done for evaluation by therapists [9].

Zannatha *et. al.* [10] proposed a stroke rehabilitation system for upper limbs using Kinect, an interactive virtual environment, a humanoid robot and devices producing ergonometric signals. The system architecture they created allows to integrate monitoring programs to evaluate the progress of a patient by a human rather than an automated evaluator.

Robotic systems are also used for early detection and treatment of Autism Spectrum Disorder (ASD). Ranatunga *et. al.* [11] used a humanoid robot to perform interactive upper body gestures which the child can imitate; the child's motion was then recorded using a motion capture system. Similarity of performances of the child and robot was measured using Dynamic Time Warping. Greczek *et. al.* [12] performed a study that examined the effects of a humanoid robot giving the minimum required feedback – graded cueing – during a one-on-one imitation game played children with ASD. The motion evaluation was done with Kinect and in successful imitation cases no prompts were given to the child. Their study suggests that varied feedback may be more effective, and less frustrating, than fully descriptive feedback in a child-robot imitation game.

McCarthy *et. al.* [13] showed the use of social robots as an intervening therapeutic aid in the rehabilitation of paediatric patients. In this preliminary study they implemented lower body rehabilitation exercises and associated voice commands in Nao but the system does not include the monitoring aid yet. Tsiakas *et. al.* [14] proposed an interactive learning and adaptation framework with a use case in robot assisted therapy, presenting a robot yoga trainer that monitors a yoga training session. By adjusting the session parameters based on visual activity evaluation through depth data, the robot assists the user to complete the session.

THERAPIST, a socially interactive robot for neuro-rehabilitation assistance is proposed by Calderita *et. al.* [15]. Initial experimental studies showed that pediatric patients can be easily driven into highly attentive and collaborating attitudes by letting them interact with a robot. However, in order to be safe and robust, this robot was teleoperated, requiring considerable supervision effort from clinic professionals.

Martin *et. al.* [16] presented NAOTherapist, a system that allows the therapist to design a complete therapy to be carried out by the robot. NAOTherapist allows the therapist to start each session with the robot, to evaluate the patient condition over the therapy and to generate reports at the end of a session. The system is well designed in terms of providing therapist a tool to select motions and design a therapy, the system includes Kinect to measure the performance of the child and to give positive feedback only eye leds were used and for improving the performance verbal and physical feedback was used.

Robots in the role of a coach, providing motivation for healthy living, diet and exercise is another area where SAR systems are used. This approach has been studied by several research groups. A pilot study by Ramgoolam *et. al.* [17] compares the effects of coaching delivered by a social and mobile humanoid robot health coach versus a human health coach on young adults. Their results suggest that the robot coach was as successful as the human coach at socially and motivationally engaging participants during the workouts. In this study motion tracking was not included and they expect to increase the effectiveness of the health coach robot with this improvement in future studies.

Matsusaka *et. al.* [18] developed an exercise demonstrator robot, TAIZO, to aid human demonstrators during simple arm exercises with a training group. However, this robot was not autonomous, as it was controlled via key input or voice by the lead human demonstrator, and it did not have sensors by which to perceive the users.

Park *et. al.* [19] presented the effectiveness of robot social skills for enhancing the social interaction in physical training by conducting a controlled experiment with 28 participants using the NAO robot. They showed there were significant differences in naturalness, perceived intelligence, appropriate gaze and sustainability between the control group where social skills were not used, and the experimental group where social skills were used by the robot.

There have also been several approaches aimed at using robots to address obesity and provide advice or coaching to increase physical activity. Cory Kidd and Cynthia Breazeal have explored in-depth robotic solutions to weight problems [20, 21]. Their robotic weight loss coach is compared to a standalone computer and a paper log in a controlled study. Results show that participants track their calorie consumption and exercise for nearly twice as long when using the robot than with the other methods and develop a closer relationship with the robot. Both are indicators of longer-term success at weight loss and maintenance and show the effectiveness of sociable robots for long-term HRI. However, the robot did not directly participate in the physical activities.

Fasola and Matarić [22] proposed a novel SAR system where robots act as coaches. Their agent is both an administrator and an active participant in the health-related activity, resulting in a unique characteristic of the system: The social interaction between the robot and user is not only useful for maintaining user engagement and influencing intrinsic motivation, but is also necessary to achieve the physical exercise task. Their work developed a set of design principles and robot qualities for success that can be summarised as motivating, highly interactive, personable, intelligent and task-driven. The principles were derived from, and evaluated on, a range of workout and memory games.

Another research area in robotic rehabilitation is mimicking human motions to learn desired exercise patterns from professionals. Howard *et. al.* [23] use a depth sensor to capture human exercise movements, which can then be translated to the robot's joint configuration space. This initial work focuses on motion translation where the human end effector (i.e. hands and feet) poses are tracked and then applied to the robot. Another study proposed a robotic fitness coach that learns a set of physical exercises from a professional trainer, and assists elderly subjects in performing these gestures [24]. However they do not assess interaction aspects such as perceived usefulness and perceived ease of use.

Ros *et. al.* [25] implemented a robotic system capable of autonomously instructing dance sequences to children while displaying basic social cues to engage them in the task. The decision making and the overall behavior management of the robot is completely autonomous but the vision and speech components are wizarded.

A prototype telepresence robot combined with a NAO upper body is proposed by Wilk *et. al.* [26] to facilitate remote communication between the patient and the clinician and to complete supervisory exercise coaching. The Nao portion of the demonstration was strictly an automated program, with little feedback: there was no two-way communication. The questions regarding the participants' interest level in the robot, their willingness to recommend the robot to friends, and their willingness to exercise with the robot again were used to describe the participants' enjoyment while working with the robot.

In this thesis we present our system design methodology, and implementation details of our socially assistive child-robot interaction system that aims to motivate children to engage in physical exercise. Our SAR system provides active guidance and online feedback through monitoring the performance of the child via wearable IMUs and does automated evaluation rather than using the input of an external human evaluator as in [10,13,15,17,18]. In our system we also focus on both engagement during exercise and completion of the motions by correcting them through feedback instead of only showing the motion. Our robotic exercise coach is both an administrator and an active participant in the exercise session as in [22], which makes the social interaction useful for both maintaining user engagement and completing the physical exercise task. We aim to apply a number of previously validated design principles in our proposed system while focusing on children instead of the elderly or adults, which are the target subjects of most of the previous studies.

The rest of the thesis is organized as follows: In Chapter 3 we explain the approach and methodology of our system design. In Chapter 4 , we discuss preliminary offline versions of our evaluation system and preliminary experiments and studies with the guidance of physiotherapists in exercise theraphy for upper arm rehabilitation of children. In Chapter 4, Section 4.1 presents the offline evaluation system design and experiments with Kinect and Nao, Section 4.2 presents offline analysis and preliminary study of system with Exl-S3 Inertial Measurement Units and Nao. In Chapter 5, we describe our online SAR system architecture, implementation details of motion evaluation and feedback selection, interaction design and preliminary experiments. Chapter 6 presents the experimental design and the results of our study with 19 children and detailed system evaluation. We conclude the thesis with Chapter 7 with discussion and future work.

3. APPROACH AND METHODOLOGY

The main goal of this research effort is to contribute to the evolution of long-term child-robot interaction in an exercise scenario. In order to implement an adaptive robot behavior that catches the child's attention, the system should have the ability to motivate the child to perform the exercise motions and maintain this engagement by building trust in the task-oriented interaction.

In order to build trust in interaction, the robot should first of all be able to perform the desired actions. The proposed system includes a humanoid robot Nao that can perform numerous exercise activities; it is a programmable, 57cm tall humanoid robot with 25 degrees of freedom (DOF) and its schematic diagram can be seen in Figure 3.1. We chose to use such an embodied robot since Wainer *et. al.* demonstrated that embodied robots were seen as more helpful, watchful, and enjoyable when compared to remote tele-present robots and simulated robots [27].

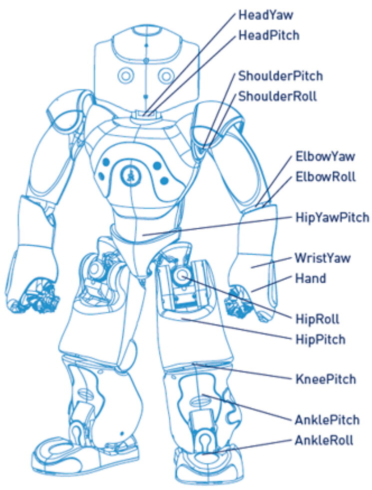

Figure 3.1. Humanoid robot Nao with its DOF annotated

In increasing the child's trust and the perception of robot's usefulness in helping to complete the desired motions, accuracy of the activity recognition procedure has a crucial role. If the user were to perceive the robot as slow or ineffective in evaluating performance, this could lead to a decrease in the user's trust and perception of the robot's intelligence [22] which may affect engagement negatively. Therefore, in order to provide a proper evaluation via activity recognition we first designed an evaluation system that mimicks the grading process of physiotherapist within a rehabilitation exercise scenario [28]. In activity sensoring, both Kinect and Inertial Measurement Units were used and tested with children. The design and evaluation details of these offline systems are presented in Chapter 4.

Csikszentmihalyi proposed that for any task to achieve a state of flow—or maximal enjoyment—in the user, it must establish a clear set of goals, combined with immediate and appropriate feedback [22, 29]. Since the primary goal of our system approach is to provide a fluid interaction to maintain user engagement in the task, after perceiving and evaluating the user's activity the robot should provide active feedback and guidance in real time.

In order to give proper feedback to the child we designed an online/real-time system that includes motivational and improving feedback. Since the potential of the system to encourage such exercise activities can be employed beyond the rehabilitation field, we designed our real-time system (with the experience gained from our previous offline studies) to behave as an exercise coach that can motivate children to improve their health and become physically active. We tested the system with an exercise coach and children during the design process. The methodology and design behind this socially assistive child-robot interaction system is explained thoroughly in Chapter 5. Figure 3.2 shows both offline and online system designs.

An experimental study was conducted with 19 healthy children to test the feasibility and effectiveness of the real-time system across a variety of user performance and evaluation measures. To measure the user's perception of robot and the interaction, we used a questionnaire which includes questions regarding their evaluation of the interaction, health coaching skills, perceived intelligence and social presence of the robot; the results of this study is presented in Chapter 6.

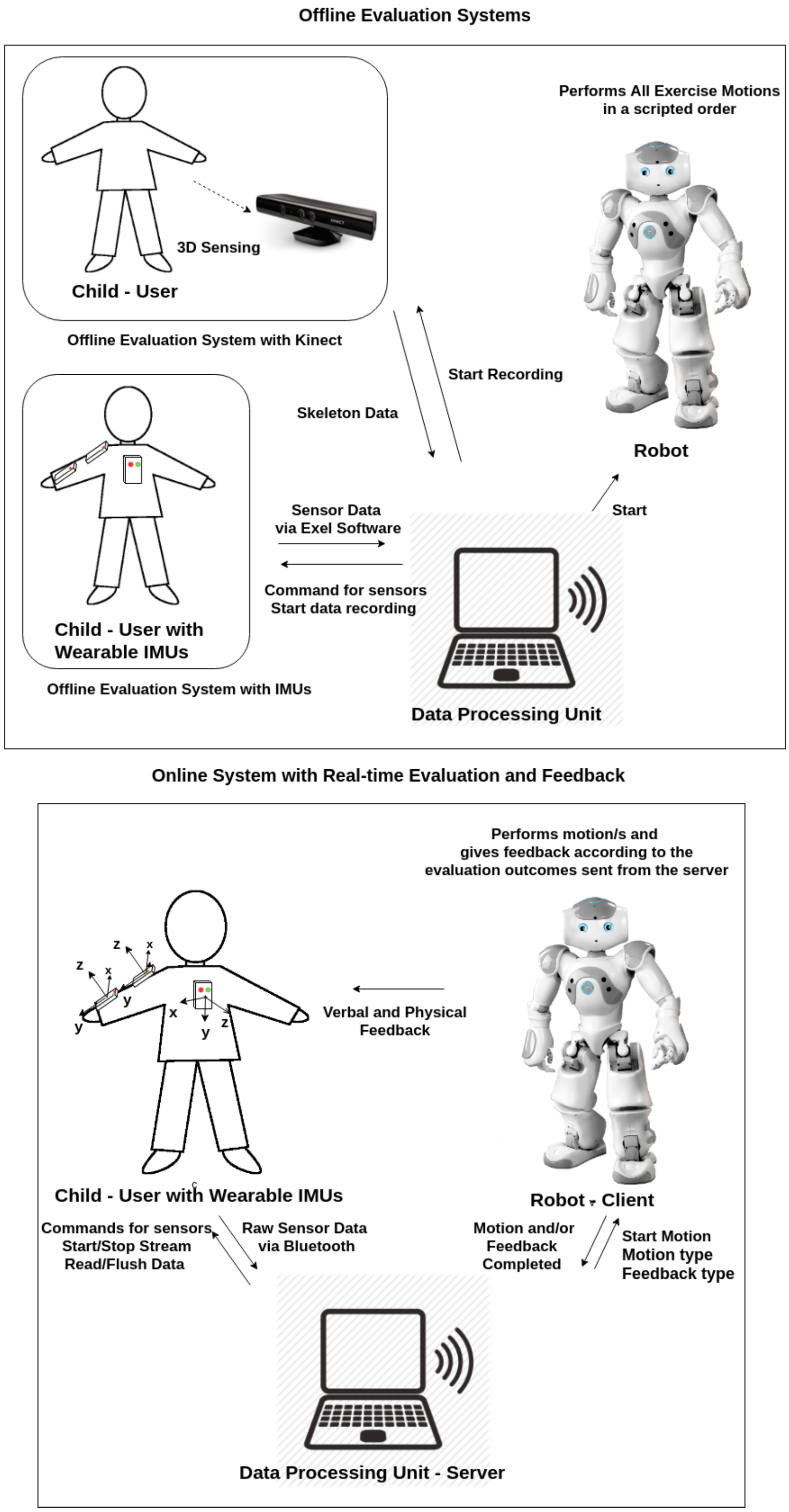

Figure 3.2. Online and offline system designs; from up to buttom: offline evaluation system with Kinect and offline evaluation system with IMUs, online evaluation and feedback system

4. OFFLINE EVALUATION SYSTEM

Our system design has evolved significantly throughout the thesis study. First we implemented an offline evaluation system for selected upper arm motions in an imitation game scenario with Nao and Kinect. However, a limitation related to Kinect vision was detected in therapy sessions: the therapist stays around the child to give some physical objects to him/her for exercises, partially or fully blocking the sensor's vision. Therefore we used weareble sensors as an alternative method. Texas Instruments' SensorTag devices were used to collect data and we implemented a simple Android application for the data collection process. However, the device was not useful to implement as an online system since the sampling frequency is not high enough for our study and it is dependent on another non-free platform to implement code for firmware which gives a very limited code size for free version. For this reason, we switched to Exls IMUs [30] for motion tracking and designed a system for upper arm rehabilitation [31].

4.1. Offline Evaluation System with Kinect

We first designed an automated offline procedure to evaluate the motions of a child who imitates the actions of a robot by using a discrete measure for similarity with an expected motion. After automatically segmenting the motions of the child Dynamic Time Warping (DTW) [32] is applied to each motion sequence in order to measure similarity with the expected motion. Evaluation is done by combining the DTW similarity and recall measure which is calculated by performed angle values during the motion. According to the therapists' feedback we did not take the latency of the motion into account while calculating the grades since it may be an indicator of unrelated factors such as attention deficit, boredom or lack of perception. Velocity and symmetry of the motion are also not considered since the main goal on these therapies is to increase the functionality of the limb which is completing the motion properly in full range. Evaluations made by our system were compared with gradings of three physiotherapists and two intern physiotherapists via Intraclass Correlation Coefficient by calculating average agreement values. Results show that evaluation of the first motion is similar to the evaluation of physiotherapist where it slightly differs in other motions. The main reason for this is that the second and third motion performances of the children are graded generously by physiotherapists since they are complex and are misperceived by children.

4.1.1. Offline Experiments

Our offline system experiments are conducted with 8 healthy children, aged between 3 to 11 where the subjects imitated three different motions performed by the Nao involving only upper-body limbs. Each of the three actions were performed by Nao three times in a repetitive manner. As the imitation scenario was being carried out, the motions were recorded by a Kinect sensor that was placed above and behind the robot, directed towards the child and recording depth and RGB data. An example of the Kinect sensor's view can be seen in Figure 4.1.

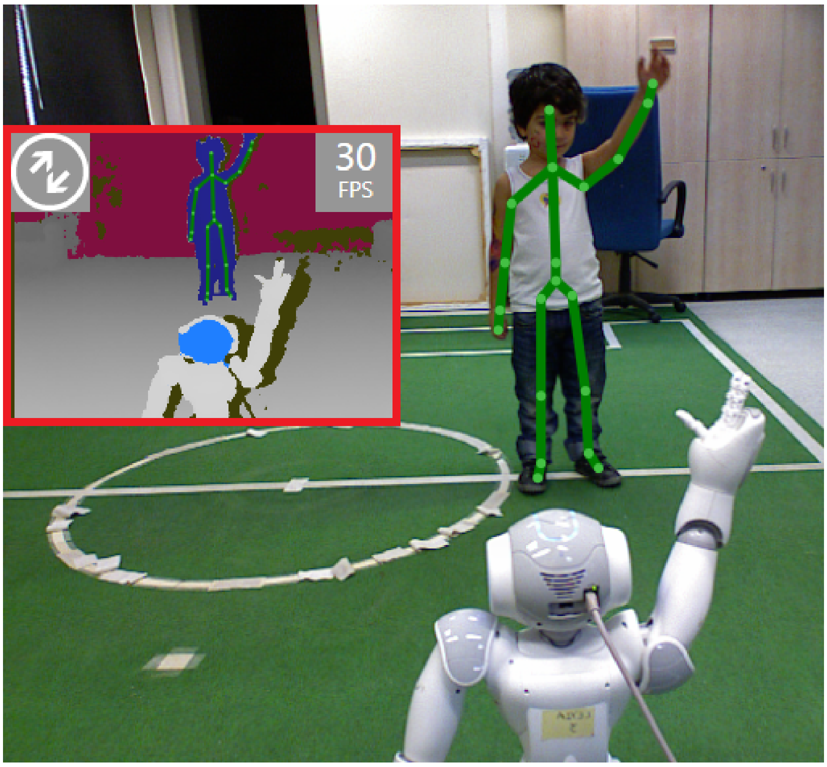

Figure 4.1. Kinect sensor's view during the offline experiments

4.1.2. Ranges and Types of Motions

The performed motions are shoulder abduction & adduction, shoulder vertical flexion & extension and elbow flexion (see Figure 4.2). Human movements are described in three dimensions based on three planes of motion that pass through the human body, namely sagittal, frontal and transverse planes [33]. The selected motions are explained briefly below:

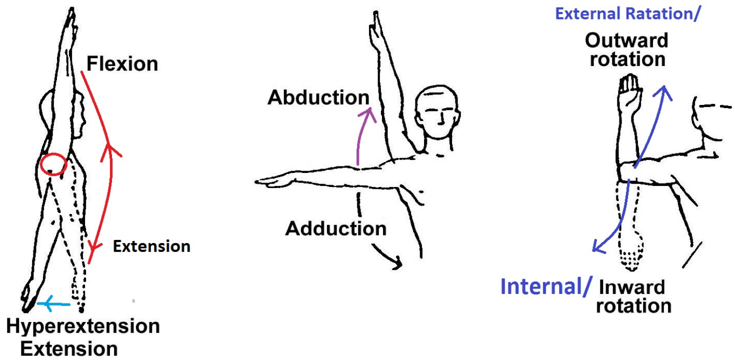

Figure 4.2. Types of motions in offline study; from left to right: shoulder vertical flexion & extension, shoulder abduction & adduction and elbow flexion

- Shoulder Abduction & Adduction: Bringing the arms up sideways on transverse plane.
- Shoulder Vertical Flexion & Extension: Raising/lowering the arms up/down straight forward/backward on the sagital plane.
- Elbow Flexion: Bringing the lower arms to biceps on both sagital and transverse planes.

Selected motions were manually created and exported using the Choregraphe [34] software. Each individual move is represented by a set of joints along with sequences of control points at given timesteps. A smooth motion is obtained through an interpolation available in Nao software. Selected poses from Nao's motions can be seen in Figure 4.3.

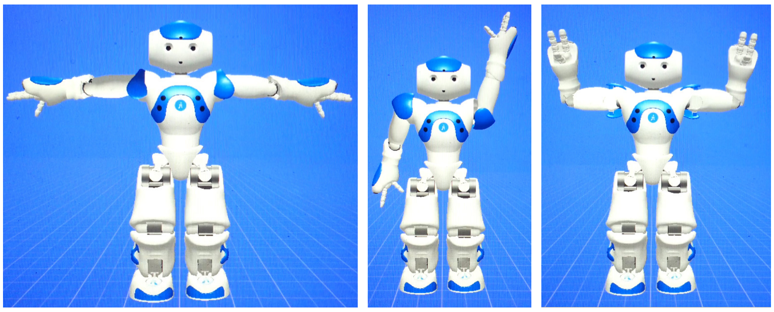

Figure 4.3. Selected poses of Nao's motions

4.1.3. Ground Truth Evaluations

The recorded RGB videos of the children were watched and evaluated by three physiotherapists and two intern physiotherapists that work in a rehabilitation facility. Each of the three motions were performed by an adult and with the approval of the therapists used as the baseline motion to be compared with the children performances. Each motion of the children was rated on a five-level Likert scale, where 1 represents a very weak imitation and 5 represents a very strong imitation. Angular data and the plane of the motion were the main criteria for the evaluation process.

4.1.4. Offline Data Processing and Motion Segmentation

In order to eliminate small angle fluctuations, we smoothed the data using moving average with a window size of 50 frames as a preprocessing step.

Our data are composed of motions such that the initial pose of each motion is the same as the last pose. Hence, we are expecting a motion to consist of three peaks that creates a curve, that form a sequence of either minimum-maximum-minimum or maximum-minimum-maximum.

The whole motion data of a child are segmented into sub-motions using an automated procedure: the local peak values are detected from the smoothed data and the peak values are then aligned to the unsmoothed data to find the real peak points.

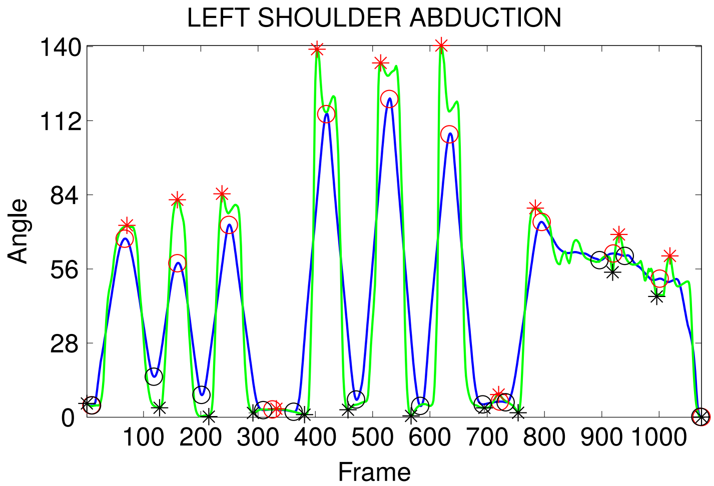

Figure 4.4. Peak detection and motion segmentation

An example of this procedure can be seen in Figure 4.4, where the circles are maximum/minimum peak points on the smoothed data while the stars are aligned versions of them on the unproccessed data.

In order to compare each motion of the child with a baseline motion DTW was implemented using a window size of three frames (detailed information about DTW can be found in next section). The sequences of two motions are warped nonlinearly in the time dimension to determine a measure of their similarity. The change between warped and initial sequences are then analyzed for similarity as explained in the next Section.

4.1.5. Evaluation of Motions with Similarity and Recall Measure

Similarity of a base motion sequence and some segmented motion data is calculated using two different measures: a similarity error calculated via DTW and a recall metric calculated from the number of angles matched between the range of motions. DTW is an algorithm for measuring similarity between two temporal sequences which may vary in speed [32, 35] . In an exercise scenario the range of motion is the main critearia for the evaluation and the speed is not considered. Therefore, we compared the child's motion with the baseline motion via DTW algorithm shown in Figure 4.5.

Get signal 1: $s = array[1..n]$

Get signal 2: $t = array[1..m]$

Determine window size: $w = int$

Create distance array: $DTW = array[0..n, 0..m]$

Adapt window size $w = max(w, abs(n - m))$

Initialize distance array

for $i = 1$ to n **do**

  for $j = 1$ to m **do**

    $DTW[i, j] = infinity$

  end for

end for

$DTW[0, 0] = 0$

Calculate distance value dynamically

for $i = 1$ to n **do**

  for $j = max(1, i - w)$ to $min(m, i + w)$ **do**

    $cost = d(s[i], t[j])$ **where** $d(x, y) = \|x - y\|$

    $DTW[i, j] = cost + minimum(DTW[i-1, j], DTW[i, j-1], DTW[i-1, j-1])$

  end for

end for

return $DTW[n, m]$

Figure 4.5. Pseudocode of DTW algorithm that calculates similarity distance between two temporal sequences which may vary in speed

Frame by frame similarity values of each aligned motion are first calculated via DTW. These are then recalculated by penalizing the larger differences more, while tolerating differences within 10% of the expected angles of the corresponding baseline motion. The penalty function is displayed in Figure 4.6.

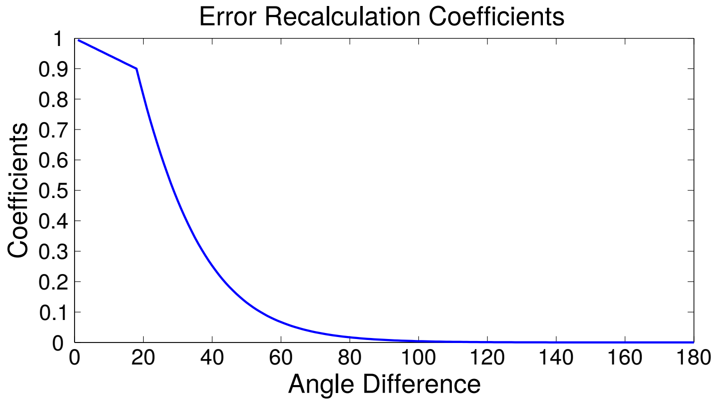

Figure 4.6. Penalty function of similarity; larger error is penalized exponentially more while up to 10% error is tolerated

According to the feedback of the physiotherapists who we worked with, the angular range that is performed by the user is an important measure in physiotherapy for functional disabilities. Taking this factor into consideration, angles that are performed within the baseline motion and child's motion are extracted. The possible range of angles is determined for relevant joints in their own ranges as explained in Section 4.1.2. Each angle is classified as *"in range"* or *"not in range"* by assigning 1 and 0 values respectively. Accuracy, precision, recall and fall-out measures are calculated. For this study, we only used the recall measure in our evaluation process. The recall measure represents how much angular area covered by the baseline motion is also covered by the child's motion in the selected frame window. It is calculated by the equation 4.1 where

- True Positive: Number of angles that are in range of the baseline motion are also in range of the child's selected area of motion.
- False Negative: Number of angles that are in range of the baseline motion but are not in range of the child's selected area of motion.

$$Recall = \frac{TP}{TP + FN} \tag{4.1}$$

DTW similarity error is expected to be low, whereas the recall measure is expected to be high for high similarity in two motions. A grade from 0 to 1 is automatically given by the system and then converted to 5 point scale in order to be comparable to the physiotherapists' evaluations. In the grading phase, only the main joint that is most related to the expected motion is considered. Similarity calculated for the right arm performance and the left arm performance are averaged.

The baseline data include three trials for each action. Each trial of the child is compared seperately with these three motions. The motion of the child is graded by choosing his/her best performance since the functionality of the arm is the main objective in upper-arm rehabilitation. In Figure 4.7 the first row shows the similarity of each automatically segmented motion of the child to the three baseline motions in each rows. Base Motion 1, Motion 2 and Motion 3 represent the three trials of the shoulder abduction motion of the baseline data. Similarities are represented via gray scale where white and black color means maximum and minimum similarity respectively. Second row shows the mean similarity over three comparisons for each segmented motion of the child. The last two rows show processed right and left joint data of the child correspondingly.

4.1.6. Comparing System's Grades with Physiotherapists' Grades

Evaluation grading is compared with physiotherapists' evaluations by calculating the average agreement values via Intraclass Correlation Coefficient (ICC) which is a measure of the reliability of ratings [36]. In order to compare our evaluation results with physiotherapist evaluations we calculated the agreement value between all raters. Calculated agreement values can be seen in Table 4.1. In each run we leave one evaluator out and calculate the average agreement value (e.g. Physiotherapist1 out, Intern1 out). The ICC algorithm is also run with all raters including our proposed system

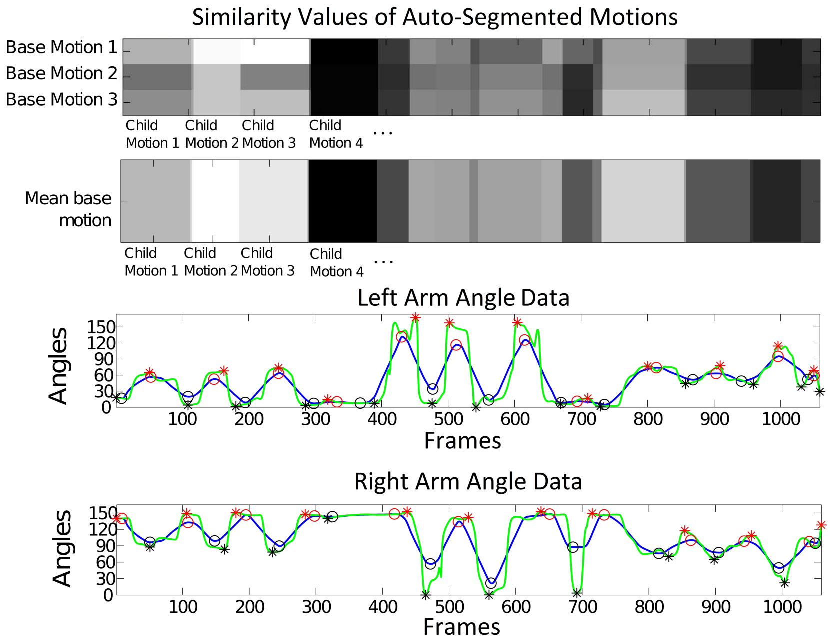

Figure 4.7. Similarity values of auto-segmented motions of the 4th subject with three baseline shoulder abduction motions

which is the "All included" case in Table 4.1.

It can be concluded that our proposed evaluation system increases the agreement for the first motion, whereas this does not hold for the other motion evaluations. For the second and the third motion evaluations, agreement values differ slightly when our system enters the calculations. The reason for the high agreement for the first motion grades is that the main and only criterion of evaluation of the first motion by the physiotherapists is one joint angle that we also considered in our proposed system. On the other hand, the second and third motion evaluations of physiotherapists include more joint values than we considered. Since we did not combine all joint values while calculating the similarity, the agreement between other evaluators and our system decreased slightly. Another reason for this difference is that the second and third motions are harder and more complicated so that they were likely misperceived by children; during the evaluations therapists compensated this misunderstanding by giving higher grades to the children while our proposed system considered the angles only.

Table 4.1. Agreement values of different combinations of evaluators for rating of each motion

Rater	Motion1	Motion2	Motion3
All included	0.7935	0.8961	0.952
Intern1 out	0.8193	0.8669	0.9344
Intern2 out	0.7148	0.8853	0.9357
Physiotherapist1 out	0.7334	0.8578	0.9384
Physiotherapist2 out	0.7725	0.8675	0.9417
Physiotherapist3 out	0.7666	0.8853	0.9356
System out	0.759	0.9003	0.9665

4.2. Offline Evaluation System with IMUs

Up to this point in our study, we used the Kinect sensor to detect angles of motions and we observed a limitation during experiments: The therapist stays in front of the children and gives them an object such as ball and pen during a functional activity which creates a barrier to Kinect vision. Therefore, we decided to use wireless inertial measurement units (IMUs) as an alternative way to monitor motions in real-time. We use the EXL-S3 wireless IMU modules from EXEL [30] (see Figure 4.8). The size of the module is 54mm $\times$ 33mm $\times$ 14mm and it weighs 22g. It features tri-axial accelerometer, gyroscope and magnetometer; it outputs its orientation in quaternion form where the estimation is improved with embedded Kalman filtering. EXEL IMU Controller software is used for data collection and each data sequence is visualized during the experiments; an example can be seen in Figure 4.9.

4.2.1. Activity Selection

The results from our previous study showed that physiotherapy motions known from literature (as in Figure 4.10 and 4.11) can be misunderstood by the child. Such misunderstandings affect how the child performs the actions. In order to make it easier

Figure 4.8. EXL-S3 Wireless Inertial Measurement Unit

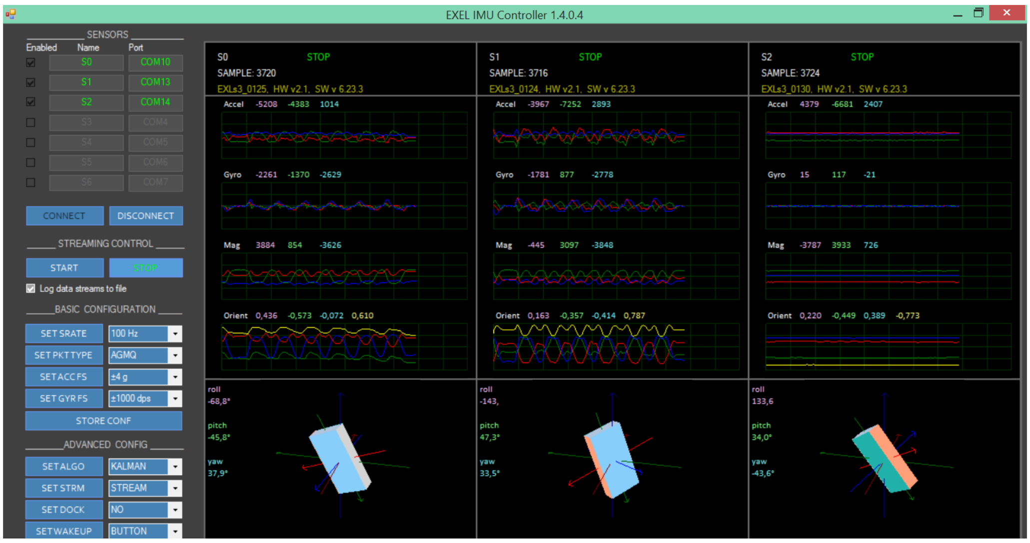

Figure 4.9. EXEL IMU Controller software interface

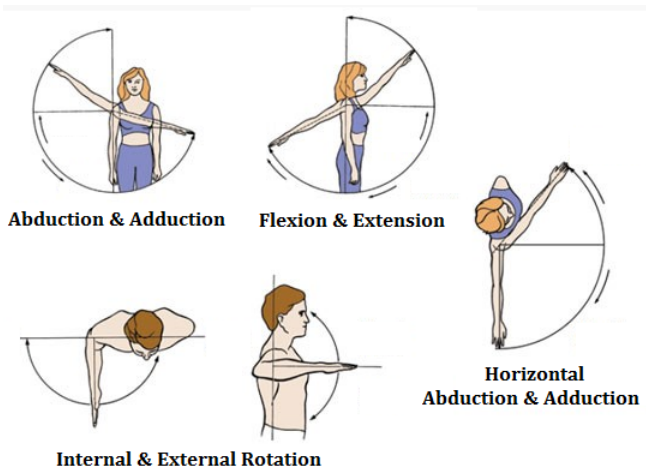

Figure 4.10. Physiotherapy motions of the arm

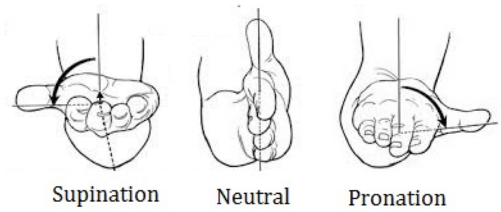

Figure 4.11. Physiotherapy motions of the lower arm

for the child to understand these, functional activities that include the desired therapy motions were selected together with the therapist. These activities are as follows:

- Opening a door with a key: It includes supination and pronation motions and the person rotates his/her hand like holding and turning a key.
- Touching the opposite shoulder with hand: It includes horizontal adduction and the person touches his/her left shoulder with the right arm without rotating the body.
- Taking an object from back of the neck: It includes external rotation and the person takes an imaginary object, from the back of his/her neck.
- Taking an object from the back: It includes internal rotation and the person takes an imaginary object, from his/her back.
- Reaching an object above the head: It includes vertical shoulder flexion and the person tries to reach an imaginary object, above his/her head.

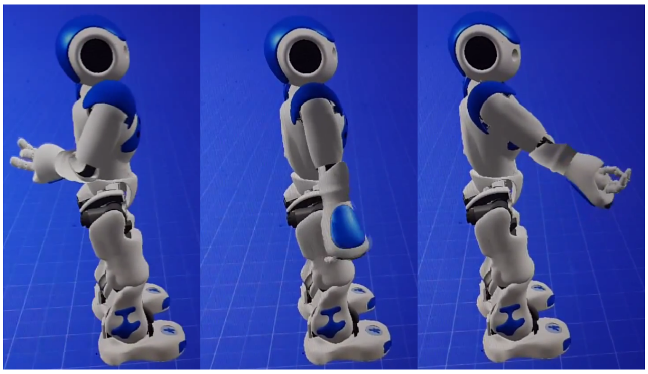

Figure 4.12. “Taking a pen from back action” of Nao

Selected functional actions were implemented on Nao via Choreographe. First, therapists were consulted and the simulated actions, such as the one shown in Figure 4.12, were validated. Then, during the experiments, these pre-programmed actions were played.

4.2.2. First Meeting with Nao and Ethical Concerns

In order to observe how children react to the Nao robot and to the wearable IMU modules, we introduced the system in a rehabilitation center for hemiparetic and diparetic children. Following ethical principals were applied:

- Proximity to the child shall not be in personal space at first but child can come closer if he/she desires
- Movements shall be smooth in order not to harm the child
- Children shall be asked if they would like to play with Nao or not

With these considerations, one 11 years old girl and one 5 years old girl experienced the system. An example scene can be seen in Figure 4.13. Both were observed to enjoy

Figure 4.13. Nao in the therapy

the robot. They wanted to do movements together and wanted the robot to imitate themselves. The younger child wanted the sensors to be removed after 5-6 minutes but the older one used them throughout the session. We observed that a more appealing cover might be necessary to make the IMU modules more attractive for small children.

Figure 4.14. Position of sensors during experiments

4.2.3. Positioning of Sensors

In the first experiments with IMUs worn by healthy children we collected data from 2 children by positioning two sensors aligned towards the posterior frontal plane. We chose a specific distance from the shoulder to place the upper arm sensor and a specific distance from the elbow to place the lower arm sensor. After examining the data of each functional activity, visualization of the data was not clear enough to distinguish actions, especially in the supination action where the child moves only the lower arm. Therefore, the position of the lower arm sensor was changed and the final positioning is shown in Figure 4.14.

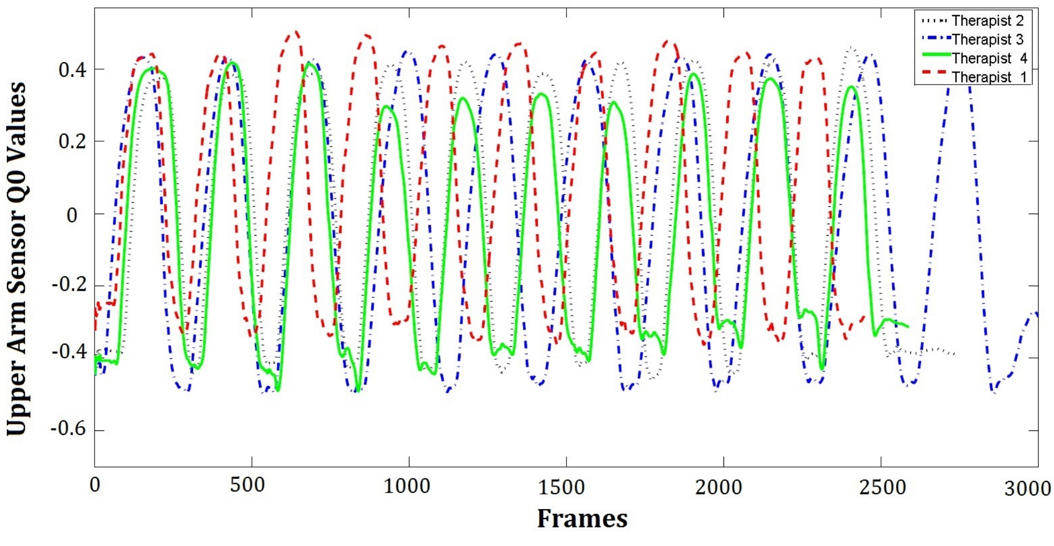

Figure 4.15. Upper arm sensor, 3rd quaternion axis values of 10 sequential reaching actions of 4 therapists

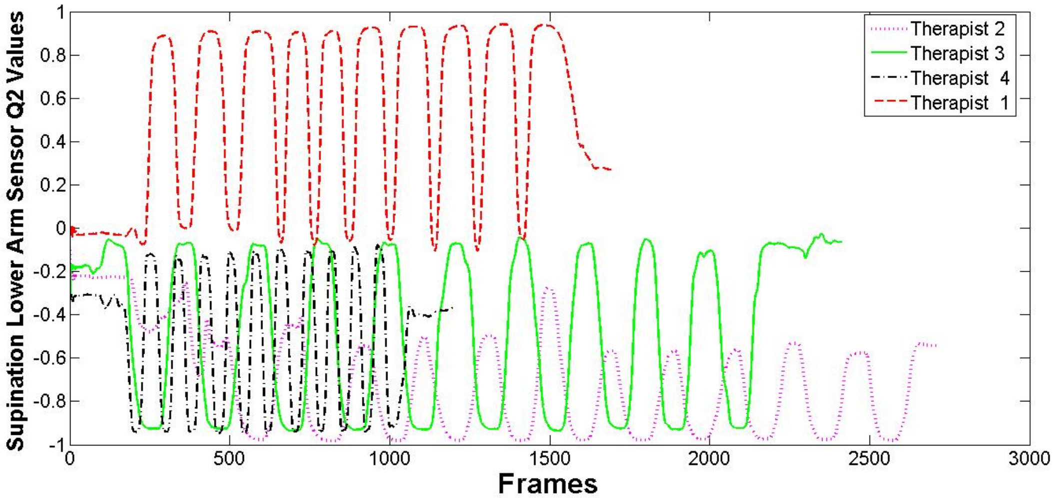

Figure 4.16. Lower arm sensor, 2nd quaternion axis values of 10 sequential supination actions of 4 therapists

4.2.4. Capturing Therapist's Motion

In addition to recording children's motion, we collected data from 4 physiotherapists. Each physiotherapist performed 5 selected functional actions 10 times and data was recorded by 2 sensors positioned as in Figure 4.14. They performed the actions in full range with proper angle values.

4.2.5. Signal-based Statistical Features

Before pre-processing, we visualized the data to ensure that the quaternion data of specific motions match as in Figure 4.15. Then basic statistical features such as mean and variance were calculated since these features provide high performance across a variety of activity recognition problems [37]. The features showed that physiotherapy actions have apparent relationship with specific quaternion axis data as outlined in the following and summarized in Table 4.2. For example, in the 'key turning' action data, we observed differences in means and found out that the initial position changes from person to person, affecting the mean (see Therapist 1 and Therapist 4 in Figure 4.16). During the experiment, Therapist 2 was observed to perform the motion with about half the amplitude (compared to other therapists). In Figure 4.16, the smaller peaks in the 2nd quaternion axis of Therapist 2 can be observed. This implies that there may be a useful relationship between standard deviation of the data and the rotation amplitude of the arm.

Table 4.2. Relationship between actions and quaternion axis data

Action	Lower arm sensor	Upper arm sensor
Key turn	2. quat. axis	-
Touch shoulder	1. quat. axis	1. quat. axis
Take from neck	2. 3. 4. quat. axes	1. quat. axis
Take from back	2. quat. axis	1. 2. quat. axes
Reach	3. 4. quat. axes	3. 4. quat. axes

4.3. Towards a Real-Time Feedback and Evaluation System

In offline studies we focused on the evaluation of exercise motions performed by the children during exercise sessions with the robot. However, in order to achieve an interactive engagement that catches the child's attention and motivates the child to imitate the robot, the robot should serve real-time feedback as well as real-time evaluation of the actions of the child, instead of offline evaluation. As a result, we shifted our focus to designing a real-time SAR system from this point on; our main motivations for this choice are given below.

During initial studies we observed many theraphy sessions of children in physiotherapy hospitals. We perceived that for each child the range of motion that he/she is able to perform is different according to the stage and type of their medical conditions. Since this range is different for each child the therapist gives positive or encouraging feedback when he/she performs the motion within his/her individual maximum limits. Similarly, the minimum limit of the child is also well-known by the therapist (which helps the therapist to understand if the child has willingness to do the motion or not) and the correction feedback is given accordingly. Another observation is that the therapist learns and knows the emotional state of the child during several sessions, selects feedback phrases according to the mood of the child. Otherwise, giving a correction feedback when a child needs an encouragement would potentially affect the success of the therapies negatively. In order to provide a feedback system for child rehabilitation all these issues should be considered and individual systems should be designed with the guidance of the therapists, which are potentially easier to integrate into an online system such as the one we are aiming to design.

Since our main objective is to evaluate the activity and give proper real-time feedback to engage interaction, we designed an online SAR system and used it in physical exercise concept with healthy children instead of rehabilitation concept which requires special or individual system for each user. We have attended numerous exercise lessons in primary schools and observed that the children do not want to attend warming exercises (which are similar to the basic rehabilitation exercises) at the beginning of the lessons and the exercise coach motivates them to do these exercises by letting them play any game they wish at the end of the lessons (if they complete the exercises). Therefore, we observed a potential need of motivation during exercise activities in primary schools where our SAR system can provide a solution.

In addition, the potential of a real-time system to encourage exercise activities can be employed beyond the rehabilitation and exercise lessons, and in the future children can have their own robot that behaves as an exercise coach that can motivate them to improve their health and become physically active.

5. REAL-TIME FEEDBACK AND EVALUATION SYSTEM

The overall system design is composed of 3 IMUs, a Nao robot and a data processing unit as in Figure 5.1. First we will explain online data acquisition (from IMUs to data proccessing unit via Bluetooth) and processing. In order to create an activity flow we designed our system step by step in order to test it in each phase. We collected data to determine thresholds for rule based evaluation methods to give feedback for straightforward motions like horizontal abduction and vertical extension. Another data collection phase is done to train our classifier and calculate Dynamic Time Warping distances for our repetitive motion *multiple vertical extension*. Thresholds, feedback and motion selections are finalized after several preliminary studies with an exercise coach and children. Details of sytem design and preliminary studies will be explained in this chapter, the experimental results of the final design will be presented in Chapter 6.

5.1. System Design

Our system is designed to be composed of 3 components (Figure 5.1): The user who interacts with the robot, the robot who interacts with the user and communicates with the data processing unit, and the data processing unit, which is composed of a laptop PC and the 3 IMUs. The robot is programmed with Python scripts and communicates with the data processing unit via a local wireless network under the role of client. The data processing unit runs a Matlab script that processes the data, communicates with the IMUs via Bluetooth serial channels and communicates with the robot under the role of server. The tasks done by the two non-user components are listed below:

- The Robot (Client)
    - (i) Waits for start command from the server

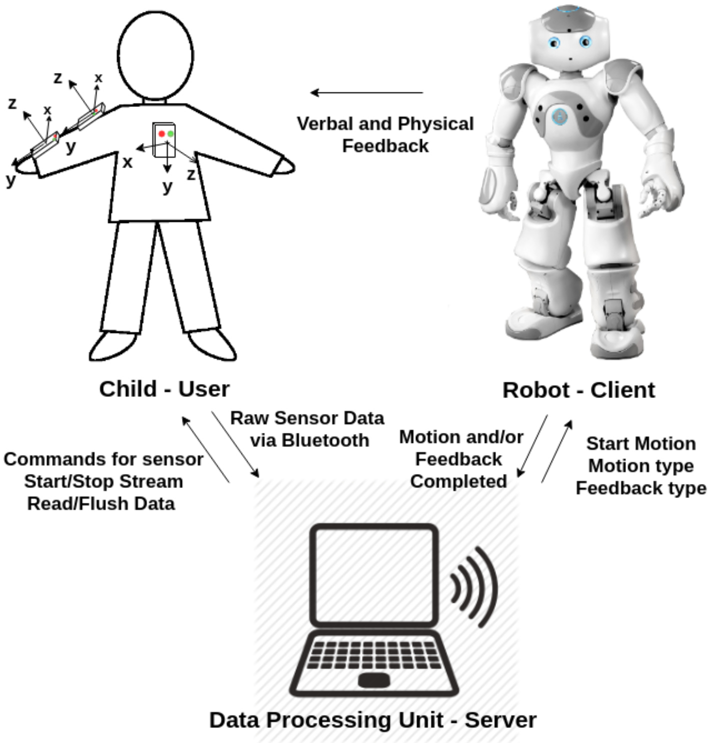

Figure 5.1. Online system design

(ii) With start command of the server starts to introduce itself

(iii) Gives information about itself conserning ethical issues and safety of children

(iv) Starts horizontal abduction motion and sends message to the server

(v) Waits for corresponding feedback message and gives positive or correction feedback of related motion to the child

(vi) Checks if the motion is changed by the server or not. If the motion type is changed, starts to do given motion type and again sends message to the server and waits for feedback type. If the motion is not changed then repeats the previous motion and sends message to the server then waits for feedaback type.

(vii) For selected motions and selected correction feedback shows the incorrect motion performed by the child back to him/herself and shows the correct motion afterwards. After completing the motion sends repetition message to the server instead of motion type message and waits for the feedback. If correction is successful and feedback is positive, gives positive feedback.

If correction is not successful and feedback is the same, again shows the incorrect motion performed by the child and then performs the right motion. This cycle continues at most 3 times in total but is not repeated if the child corrects his/her motion.

- The Data Processing Unit (Server)
  - (i) Connects to Sensors via Bluetooth serial channels
  - (ii) Sends start message to the robot
  - (iii) Waits for message of performed motion type from the client
  - (iv) When motion type is received starts to count the number of feedback and checks if it reached maximum feedback number of selected motion. If it reaches the maximum changes the motion type and sends message to the client
  - (v) Gets sensor data and calculates corresponding angles
  - (vi) According to the performed motion type, selects feedback calculation algorithm
  - (vii) Does rule based checks or classification to decide on feedback type for corresponding motion
  - (viii) Sends feedback type to the robot
  - (ix) If receives maximum number of feedback related to one motion changes the motion and sends robot the new motion type
  - (x) If receives a repetition feedback instead of the motion type, gets sensor data and gives feedback to the client without increasing feedback number according to current motion type

5.2. Online Sensor Data Acquisition and Preprocessing

In order to collect data from Exls3 IMU sensors in Matlab we used the Instrument Control Toolbox Bluetooth interface [38] which allows us to connect to devices over Bluetooth serial channels and to transmit and receive ASCII and binary data. We identify each sensor and establish two-way connection with each device. We send the appropriate command over Bluetooth (a 0x3D character) in Matlab to start real-time streaming of sampled and processed data to the controlling PC. The sensor starts sending to the host the data sampled in real-time with the selected format (which is set through EXEL IMU Controller Interface). Likewise, we send the appropriate command over Bluetooth to stop streaming (a 0x3A character). All raw inertial measurements

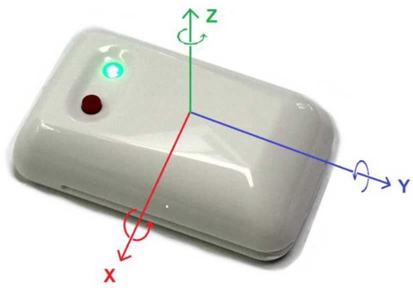

Figure 5.2. IMU module frame of reference

(linear accelerations, angular velocity and earth magnetic field) provided by the unit are with respect to the reference system shown in Figure 5.2. When in streaming mode, the EXLs3 unit can send different types of packets, depending on the value of the PACKET_TYPE parameter (set in EXEL IMU Controller Interface beforehand). With O_Type packet, the EXLs3 units sends the estimated orientation of the unit. The value of a progressive packet counter is also packed.The information includes:

- PKT_TYPE: header which identifies the type of packet.
- PKT_CNT: progressive packet number ranging from 0 to 10000.
- CHKSUM: sum modulo-256 of all previous bytes.

The orientation of the sensor is given by a quaternion (q_r, q_i, q_j, q_k) representing the rotation of the unit's body frame with respect to the earth frame. Thus when the X axis of the sensor is pointing north and Y axis is pointing west the output is the identity quaternion $(1, 0, 0, 0)$. The quaternion's components returned by the EXLs3 are normalized so that unity is represented by the value 16384 and are in 2's complement. The code in figure 5.3 is used to extract the orientation data from one packet.

Create Bluetooth channel: $sensor = Bluetooth('EXLs3\_0234', 1)$

Connect to the device: $fopen(sensor)$

Start streaming: $fprintf(sensor, '=')$ % 0x3D

Read 13 bytes which is one packet: $bytearray = fread(sensor, 13)$

Get qr value bytes: $qr = bytearray(5) + 256 * bytearray(6)$

if $qr >= 32768$ **then**

  $qr = qr - 65536$

end if

Get qi value bytes: $qi = bytearray(7) + 256 * bytearray(8)$

if $qi >= 32768$ **then**

  $qi = qi - 65536$

end if

Get qj value bytes: $qj = bytearray(9) + 256 * bytearray(10)$

if $qj >= 32768$ **then**

  $qj = qj - 65536$

end if

Get qk value bytes: $qk = bytearray(11) + 256 * bytearray(12)$

if $qk >= 32768$ **then**

  $qk = qk - 65536$

end if

$qr = qr/16384$

$qi = qi/16384$

$qj = qj/16384$

$qk = qk/16384$

Stop streaming: $fprintf(sensor, ':')$ % 0x3A

Figure 5.3. Pseudocode of the script that reads quaternion values from the Bluetooth serial channels

5.3. Sensor Positioning

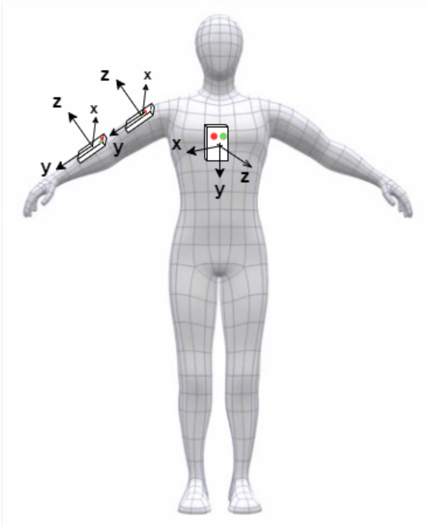

Figure 5.4. Sensor positioning

Sensors are positioned on the body as in Figure 5.4. One is placed on the chest for body tilt angle detection, one is placed on the right upper arm where z axis points to the right side of the body to calculate arm rise angle and arm vertical turn angle and one is placed on the lower arm to calculate elbow angle between upper and lower arm.

5.4. Angle Calculations from Sensor Data

We collected only quaternion values from each sensor. Quaternion-derived rotation matrix R of each sensor is calculated by the equation 5.1. From each rotation matrix; x,y,z axis angle vectors of each sensor is calculated by Equations 5.2, 5.3 and 5.4.

$$R = \begin{bmatrix} 1-2q_j^2-2q_k^2 & 2(q_iq_j-q_kq_r) & 2(q_iq_k+q_jq_r) \\ 2(q_iq_j+q_kq_r) & 1-2q_i-2q_k & 2(q_jq_k-q_iq_r) \\ 2(q_iq_k-q_jq_r) & 2(q_jq_k+q_iq_r) & 1-2q_i-2q_j \end{bmatrix} \quad (5.1)$$

$$x_{sensor} = R_{sensor} \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T \tag{5.2}$$

$$y_{sensor} = R_{sensor} \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^T \tag{5.3}$$

$$z_{sensor} = R_{sensor} \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T \tag{5.4}$$

From these, the relevant angles for our system are calculated as follows:

- Body back-front tilt angle is calculated as the angle between z axis of the world and z axis of the body sensor as in Equation 5.5.
- Body left-right tilt angle is calculated as the angle between x axis of the body sensor and z axis of the world as in Equation 5.6.
- Arm horizontal rise angle is calculated as the angle between y axis of the body sensor and y axis of the upper arm sensor as in Equation 5.7
- Arm vertical turn angle is calculated as the angle between x axis of the upper arm sensor and z axis of the world as in Equation 5.8.
- Elbow angle is calculated as the angle between y axis of upper arm sensor and y axis of lower arm sensor as in Equation 5.9.

$$\arccos\left(\frac{z_{body}^T z_{world}}{\|z_{body}\|\|z_{world}\|}\right) - \pi/2 \tag{5.5}$$

$$\arccos\left(\frac{x_{body}^T z_{world}}{\|x_{body}\|\|z_{world}\|}\right) - \pi/2 \tag{5.6}$$

$$\arccos\left(\frac{y_{body}^T y_{upperarm}}{\|y_{body}\|\|y_{upperarm}\|}\right) \tag{5.7}$$

$$-\arccos\left(\frac{x_{upperarm}^T z_{body}}{\|x_{upperarm}\|\|z_{body}\|}\right) + \pi \tag{5.8}$$

$$\arccos\left(\frac{y_{lowerarm}^T y_{upperarm}}{\|y_{lowerarm}\|\|y_{upperarm}\|}\right) \tag{5.9}$$

These angles form our derived features which are then compared to thresholds as described below.

5.5. Data Collection For Threshold Based Evaluation System

With threshold-based classification, a derived feature is simply compared to a predetermined threshold to determine whether a particular activity is being performed. This approach has been used successfully to differentiate between static postures [39]. In order to evaluate easy one posture motions and give correction or improving feedback to the child, the system needs the angle thresholds generated from real user data instead of ideal range of motion limits. As we observed in the exercise sessions in schools, children can not be expected to do motions in a perfect way like therapists or exercise coaches during the exercise.

In order to determine these real user angles, data from four children (two 9 years old girls, one 9 years old boy and one 10 years old boy) were collected doing multiple vertical extension, horizontal abduction, multiple vertical extension, and stretching motions. After calculating the angles, we determined the angle limits for each motion and correction feedback for preliminary experiments. Example calculated angle values for vertical extension motion perormend multiple times is given in Figure 5.5. Some angle limit values were changed after preliminary experiments of our system with children and exercise coach which is explained in next sections.

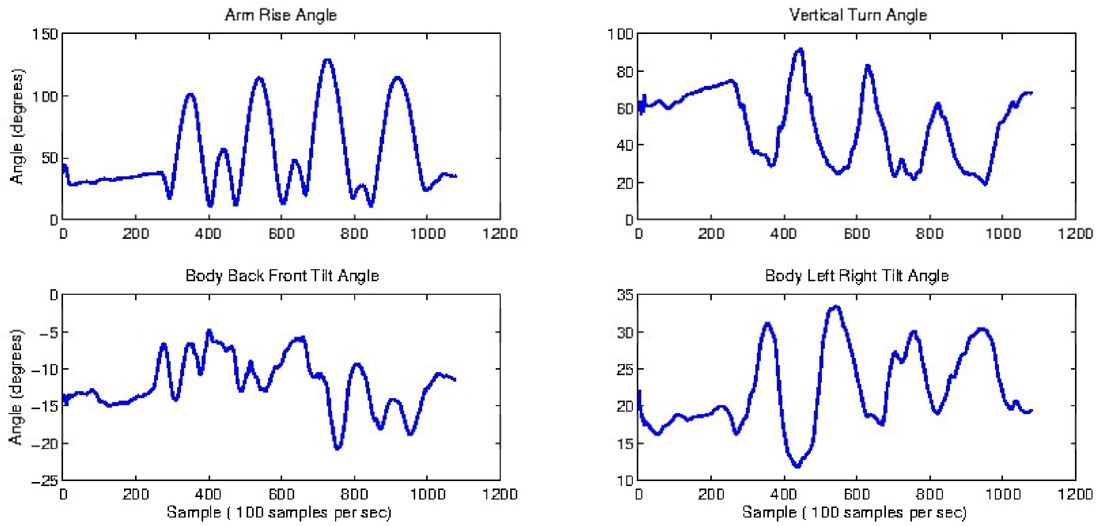

Figure 5.5. Vertical extension angles

5.6. Motion Selection and Evaluation

In order to test the threshold based evaluation system described in the previous section and test the online feedback loop, initially the online system is designed only with horizontal adduction motion with “You should not lean backward.”, “You should not lean forward.”, “Raise your arm higher.” and positive feedback with an applause sound. In a second phase, vertical extension, multiple vertical extension, stretching with one arm and stretching to two sides with both arms motions were also integrated to the action flow and tested with children again. Repetition number of each motion, length of motion performed by Nao and window size of the data collected while the child is performing the motion is determined after several preliminary tests that were done with children and an exercise coach.

5.6.1. Preliminary Tests with Children and an Exercise Coach

After initial tests with two 9 years old children and an exercise coach, it was revealed that more repetition, faster motions and fun motions like dancing was necessary according to their feedback. Furthermore, correction feedback phrases were changed and more feedback phares were added to prevent the robot from being misunderstood. Harder motions and repetitive motions were added to the system for a more active exercise session. Arm dancing is added to make the system more entertaining.

Then, the window size was determined according to the motion. For short motions (vertical extension and horizontal adduction) the processing window size was determined as 1 second (we also tried 1,5 and 2 second long windows but increasing the window size made system too slow for children). For stretching and multiple vertical extension window size was determined as 4 and 6 seconds since motions are long and the robot also performs the actions in longer periods. Figure 5.6 shows the performance of a child during second preliminary experiment.

Figure 5.6. Preliminary experiment 2

5.7. Repetitive Motion Evaluation

In preliminary experiments threshold based evaluation worked well in the evaluation of non-repetitive motions. However, it was not satisfactory in multiple vertical extension motions. In order to overcome this problem, we decided to combine DTW with an automated classifier.

5.7.1. Training Data Collection

In order to classify the repetitive multiple vertical extension motion by using the arm horizontal rise angle value we collected 108 sets of motion data (each one is 6 seconds long) from 5 subjects aged between 9 and 26. Each subject performed the motions in right and wrong ways (sometimes on purpose). Distance measure of Dynamic Time Warping is calculated by using two different baseline motions; signal based features are also calculated. In the following sections, the procedure for the multiple vertical extension motion is described, for the rest of the motions threshold based evaluation is used as described in Section 5.5.

5.7.2. Baseline Motions

In initial experiments, the exercise coach observed multiple vertical extension motions of a child and determined two baselines for minimum and maximum acceptable motions by considering the safety of the joints of children. Minimum acceptable vertical extesion angle is determined to be around 120 degrees and maximum is determined to be around 160 degrees. The ideal 180 degrees is not used since the exercise coach decided that it can harm the shoulder joints and they can normally reach that angle only after several exercise sessions. Two baseline motions and a well performed child motion can be seen in Figure 5.7.

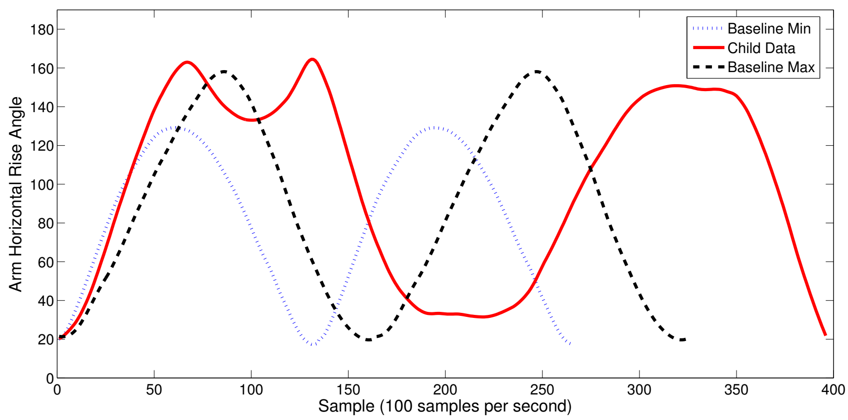

Figure 5.7. Two baseline motions and an example of well performed child motion

Table 5.1. Accuracy comparison of two classification methods.

	Naive Bayes		ClassRegTree	
Num of Class	Num of Features	Accuracy	Num of Features	Accuracy
3 Class	6	%92	6	%90
	9	%91	9	%82
4 Class	5	%86	6	%81
	8	%84	9	%81

5.7.3. Feature Selection and Classification

In order to classify performed multiple vertical extension motions, we used two distance measures coming from DTW with minimum and maximum baseline motions. Signal based features (mean, max, min and variance) are also used as features. These 6 features were used in classification. In addition to these values the difference between maximum of the signal and 120 degrees as well as 160 degrees and the difference between maximum and the minimum of the signal are also used as features in a 9 features classification case.

Classes are determined as follows:

- no motion: staying still or slight movements.
- good motion: vertical extension is done twice in the acceptable range.
- one motion: vertical extension is done only once in the acceptable range.
- wrong motion: vertical extension is not done in the acceptable range or more than two motions are done.

The latter two classes are counted as one *wrong motion class* in 3 class classification. Naive Bayes and ClassRegTree classification methods of Matlab are used with the training data. Each method is trained and tested for 3 and 4 class classification with 6 and 9 features seperately. Each combination is trained and tested 80 times by randomly shuffling the data and selecting 80% of the data as training set and 20% as testing set. In 4 class classification of NaiveBayes method, standard deviation feature was rejected by the training function of the method so 5 and 8 features were used by excluding this feature. Classification results of each method in each combination is given in table 5.1. These results led us to decide on Naive Bayes with 6 features on 3 class classification.

5.8. Activity Flow of The Robot and Feedback Design

In the activity flow; verbal feedback, LED feedback, exercise motions and generic motions are used. In order to generate salutations and generic motions we used or modified a subset of the behaviors in the NAO-Base robot behavior collection of the F.U.N. Lab at the University of Notre Dame [40]. At the beginning of the interaction robot gives information concerning following ethical issues:

- Child can think that the robot is dead or sick if it turns off so robot should mention it at the beginning that it is normal to turn off.
- Child can think that the robot is be angry with him/herself when it turns off or does not give answer to his/her questions so the robot should mention that it can not always understand him/her.
- Child can touch to the robot while it is moving and may hurt him/herself so robot warns him/her not to touch it.
- Child should know for which purpose robot is there.

Feedback phrases related to the performance of the child are divided into two groups: positive feedback and correction feedback. Positive feedback is given when the child performs the motion in the right way to praise the child. Correction feedback is given to correct or improve the performance of the child when he/she does an incomplete or wrong motion. These are also divided into two groups: repetition feedback and verbal-only feedback. In repetition feedback the robot shows a wrong motion to the child which is similar to the motion performed by the child and then shows the correct motion that should be done. In this way it can show what is wrong in the motion to the child via comparison. In verbal-only correction feedback, the robot explains what is wrong in the motion verbally.

The perceived intelligence of robots in an interaction can have a positive impact on user engagement [22]. Furthermore, the increased variability in agent behavior leads to increased engagement and self-reported desire to continue interacting with the agent [41] which means repetitive discourse tends to have a negative impact on user motivation and perceived intelligence and trust. Therefore, we tried to minimize the repetitiveness of the robot's verbal instructions and feedback by creating phrase pools for most of the positive and correction feedback types. The robot randomly selects a phrase from the corresponding feedback pool that gives the same message to the user. For instance, there are four different ways in which the robot can praise the user after performing the exercise completely (e.g., "You are super!", "Great!") and there are four different ways to give new trial notice (e.g.,"Let's try again", "Let's do it again","Let's do it one more time").

The motion flow of the robot is as follows:

- Greeting
  - Motion: Wave hand.
  - Verbal: "Hello, my name is Nao".
  - Leds: Eyes and chest leds' lights rotates with light colors.
- Information
  - Motion: Generic arm and head motions, do not understand/do motion by turning head to right and left two times slightly. Motions can be seen in Figure 5.8
  - Verbal: "I am a robot. I may not always understand what you say. My battery can be low and I can turn off. I am here to do exercise together with you. I can fall down if you touch me while I am moving. Please do not touch me. Let's do exercise together!".
- Horizontal Abduction
  - Motion: Raise arms in horizontal plane (see Figure 5.9).

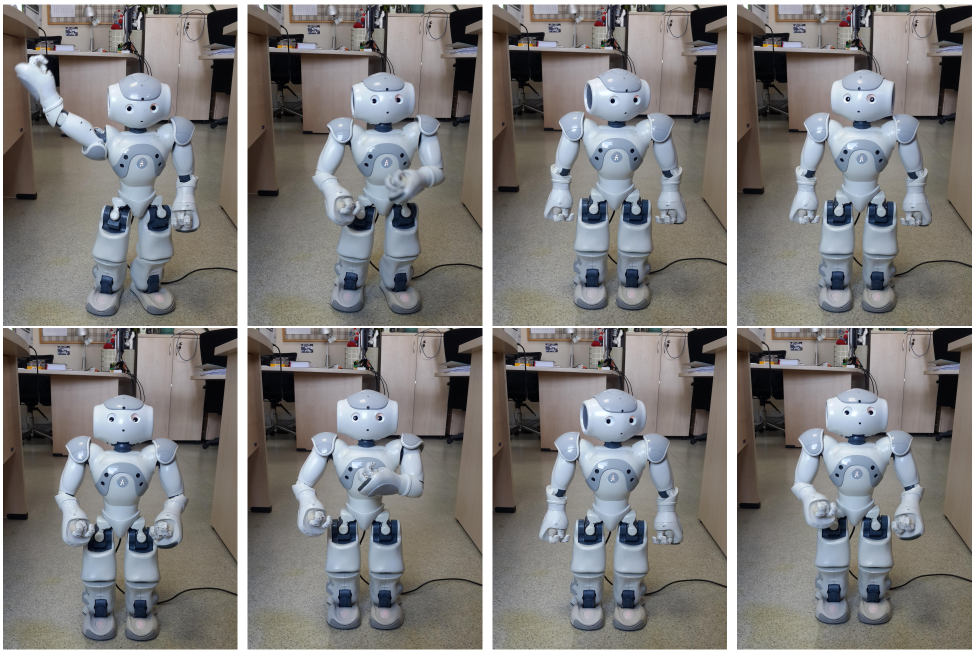

Figure 5.8. Greeting and information parts of the action flow

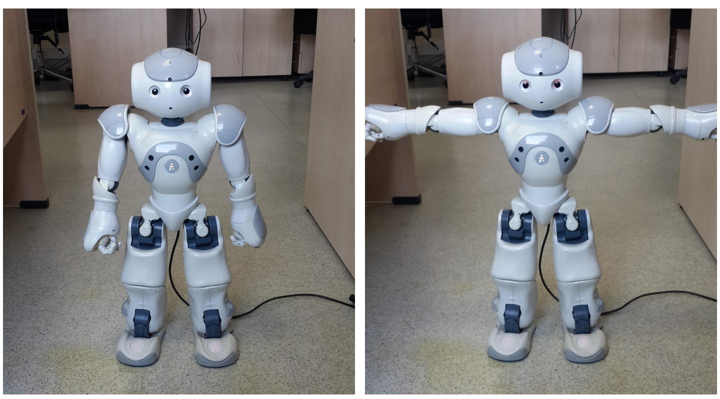

Figure 5.9. Horizontal abduction motion of Nao

  - Verbal: “Let’s raise our arms like this”.
- New Trial Notice
  - Motion: Repeated motion (previous motion).
  - Verbal: Randomly chooses one of these phrases: “Let’s try again”., “Let’s do it again”.,“Let’s do it one more time.”, “Let’s try it one more time.”
- Vertical Extension
  - Motion: Raise arms in vertical plane (see Figure 5.10).
  - Verbal: “Let’s do another exercise”.

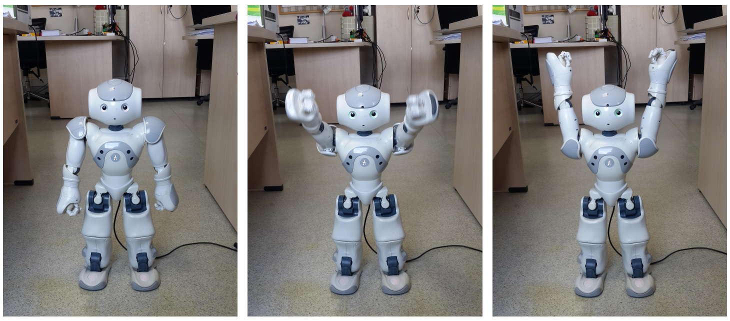

Figure 5.10. Vertical extension motion of Nao

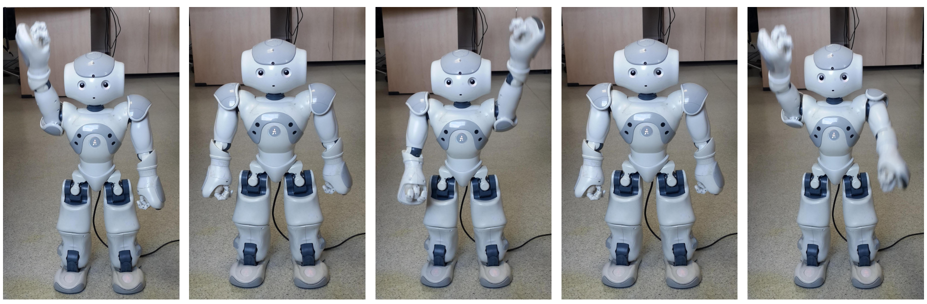

Figure 5.11. Multiple vertical extension motion of Nao

- Multiple Vertical Extension
  - Motion: Raise one arm and then the other arm two times in vertical direction (see Figure 5.11).
  - Verbal: “Let’s raise our arms one by one.”.
- Stretching
  - Motion: Raise right arm, move hand above the head and move hip to the left (see Figure 5.12).

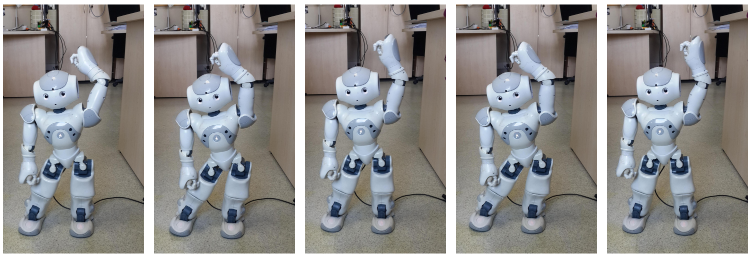

Figure 5.12. Stretching motion of Nao

  - Verbal: “Let’s raise our right arm and stretch to the left.”

- Dance
  - Motion: Twirls arms in front and to each side (see Figure 5.13).
  - Verbal: “Let’s do some dance now!”.

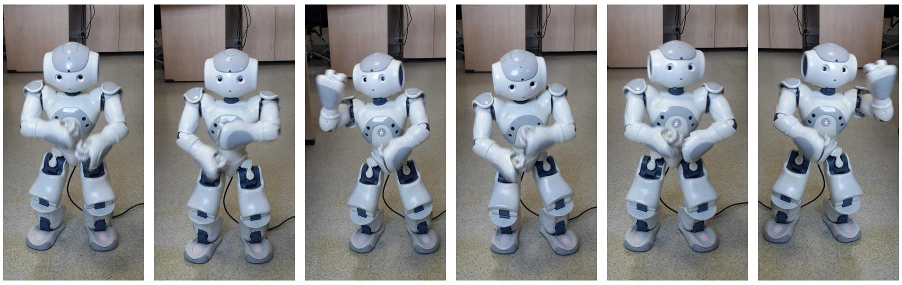

Figure 5.13. Dance performance of Nao

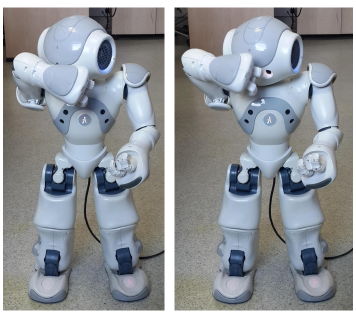

Figure 5.14. Nao is tired

- Tired
  - Motion: Moves right arm as to wipe sweat from his forehead (see Figure 5.14).
  - Verbal: “Oh I am so tired”.

- Farewell
  - Motion: Waves hand.
  - Verbal: “Thank you see you again.”.

The feedback design of the system is as follows:

- Correction Feedback
  - Verbal-Only Correction Feedback
    (i) “You should not lean backward.”,“Do not lean backward.”, “It becomes very well if you do not lean backward.”
    (ii) “You should not lean forward.”, “Do not lean forward.”, “It becomes very well if you do not lean forward.”
    (iii) “You should not twist your elbows.”
    (iv) “You should raise your arm straight two times.”
    (v) “You should tilt to the side more.”
    (vi) “You can raise your arms higher.”
    (vii) “Do not stay still. Let’s move with me!”
  - Repetition Feedback
    (i) Verbal: “You should not raise your arm like that.”. Wrong Motion: Similar incomplete motion of the child is done by holding arms not high enough around 100 degrees for vertical extension.
    (ii) Verbal: “You should not raise your too much like that.” Wrong Motion: Similar extreme motion of the child is done by holding arms too high around 170 degrees for vertical extension.
    (iii) Verbal: “You should not raise your arm like that.”. Wrong Motion: Similar incomplete motion of the child is done by holding arms not high enough around 70 degrees for horizontal adduction.
    (iv) After each type of wrong motion feedback the robot says “You should raise your arm like that” and shows the corresponding right motion.
- Positive Feedback
  - Verbal: Randomly chooses one of these phrases: “You are super!”, “You are

great!", "It is very well done."', "Great".

- Leds: All leds are on with different colors and eye leds rotate with multicolor.

6. EXPERIMENTS WITH REAL-TIME SYSTEM AND RESULTS

Experiments are conducted with 19 healthy children aged between 4 and 12. 4 children had already met with Nao in our previous offline studies and/or preliminary online experiments. Ethical issues considered in offline experments are considered also during online experiments. Before each interaction each child is asked whether they want a friend of him/her to stay in the room during the interaction. If they wanted to be accompanied by a friend we only let the ones that already finished the experiment. We explained each child that we are testing the robot and not testing their performance. Also we let them to touch the IMUs and explained what they are used for while placing them to their body. Each interaction was recorded via video. Action flow explained in Section 5.8 is run with each child at least once. However, most of the subjects wanted to repeat the whole interaction.

We evaluated the system performance via accuracy of feedback. In Section 6.1 performance related results are discussed in detail. To measure the user's perception of robots, we used a questionnaire which consisted of survey questions on a 3-point Likert scale and yes/no questions. Participants answered questions after each interaction. Questionnaire results are explained in Section 6.2.

6.1. Feedback Accuracy of The System / Performance Evaluation

The recognition system can miss, confuse, or falsely detect activities that did not occur, resulting in wrong feedback is given to the child. We evaluated the system performance by examining the accuracy of feedback of all interaction sessions during experiments. We used the following metrics to measure the performance:

- True Positive: Child completed the motion correctly and a positive rewarding feedback was given (e.g. "You are super!", "You are great!").

- True Negative: Child made an incomplete or wrong motion. In order to correct and improve the performance a proper correction feedback was given depending on the motion type (e.g."Raise your arm higher", "You should not lean backward").
- False Positive: Child made an incomplete or wrong motion and a positive feedback was given.
- False Negative: Child completed the motion correctly but a correction feedback was given or child made an incomplete or wrong motion but an unrelated correction feedback was given.

In the false negative class only 5 unrelated feedback were given to the incomplete/wrong motion. Since it is a very small set, it is included to the false negative class.

Table 6.1. Sytem feedback accuracy

True Positive	**True Negative**	Total Percantage
133	107	89%
False Positive	**False Negative**	Total Percantage
0	31	11%

The choice of the metric to be optimized depends on the application. Often it is preferable to reduce FNs at the price of FPs but in other cases, a high FP rate can make people ignore the system's notifications and eventually abandon the system [37]. We did not want our system to give too much false positive feedback since giving an encouraging feedback to a completely different motion can cause the robot to be perceived as unintelligent. As we see in Table 6.1 feedback accuracy of the system is 89% where there are no false positives detected.

6.1.1. True Positive and True Negative Feedback Analysis

Table 6.1 shows that approximately 45% of true feedback is correction feedback which means children performed incomplete or wrong motions as much as performing right motions. This result shows that there is a real need for an online correction feedback mechanism during exercise coaching besides encouraging feedback.

6.1.1.1. Motions That Get the Most Correction Feedback. When we analyse corrected motion types and given correction feedback types, we notice that the children performed wrong or incomplete multiple vertical extension and stretching more than easier motions as it can be seen in Table 6.2.

Table 6.2. Motions that get the most correction feedback

Horizontal Abd.	**Vertical Ext.**	**Multiple Vert. Ext.**	**Stretching**
14	29	32	32

6.1.1.2. Correction Feedback Types That Are Given Most. Detailed correction feedback types for each motion can be seen in Table 6.3.

Table 6.3. Correction Feedback Types That Are Given Most

Horizontal Abduction	Raise arm higher	14
Vertical Extension	Raise arm higher	28
	Make elbow straight	1
Multiple Vertical Extension	Raise arm straight two times	29
	Do not stay let's move	3
Stretching	Raise arm higher	21
	Reach side more	11

When we further analyse each feedback the following results are found:

(i) In horizontal abduction, 10 raise arm higher feedback phrases out of 14 are generated for only 3 children. This means the system could not correct their motions by showing them the wrong motion then the right motion and the children repetitively did the wrong motion. An example can be seen in Figure 6.1. On the other hand system feedback can help other children to correct their horizontal abduction motion as in Figure 6.2 and 6.3.

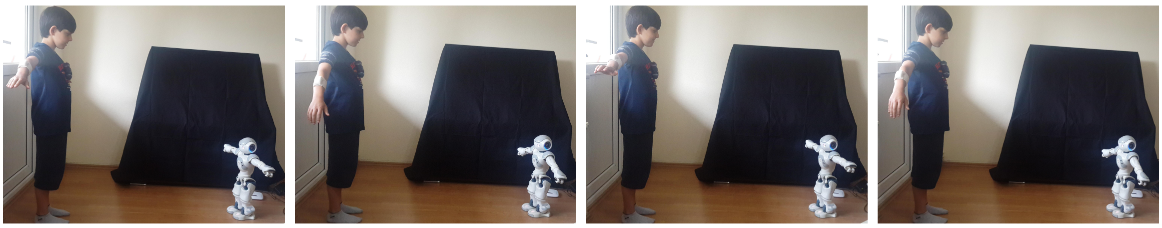

Figure 6.1. Not corrected horizontal abduction

Figure 6.2. Corrected horizontal abduction example 1

(ii) In vertical extension, 25 raise arm higher feedback phrases out of 28 are generated for only 6 children. This again means the system could not correct their vertical motion by showing them the wrong motion and then right motion and the children repetitively did the wrong motion. An example can be seen in Figure 6.4. On the other hand system feedabacks can help the same child to correct their vertical extension motion at some point as in Figure 6.5 and 6.6.

(iii) Some children who could not correct their motions in their first trial can perform the same action very well in the second trial of whole interaction as in Figure 6.7.

Figure 6.3. Corrected horizontal abduction example 2

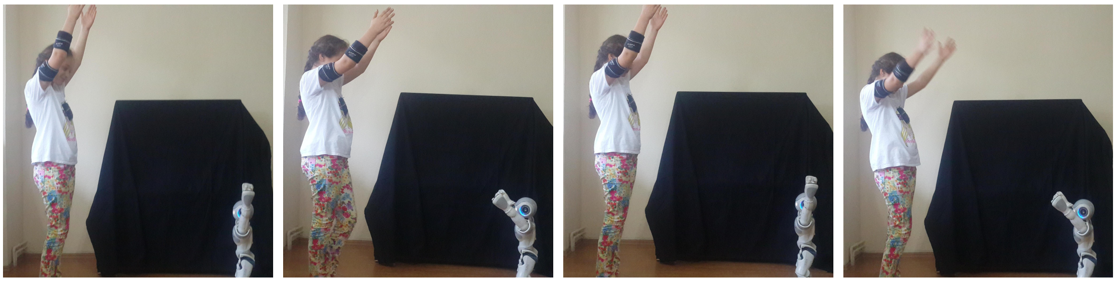

Figure 6.4. Not corrected vertical extension with *raise arm higher* feedback

This may have been possible by seeing or listening to another child's performance.

(iv) Make elbow straight correction feedback is given only once to a vertical extension and corrected motion can be seen in Figure 6.8.

(v) First two results show that some children imitate the robot without listening to the verbal feedback and by just focusing on the change in the motion angle without changing their initial wrong motion.

(vi) During experiments we observed that multiple vertical extension motion was performed mostly 2.5 times or 3 times in a right way. In other words, the child raises one arm and the other arm more than 2 times within the expected range of motion. The imitation of the robot is correct but since the motion is repetitive children could not stop just after two repetition. So in classification 2.5 or 3 times performed repetitive motions should be labeled as well performed motions instead of incorrect motions. Examples for this kind of well performed but low graded motions can be seen in Figure 6.9.

(vii) In stretching motion, robot cannot perform the motion exactly the same as a human since it cannot move its chest. Instead of moving the chest in stretching,

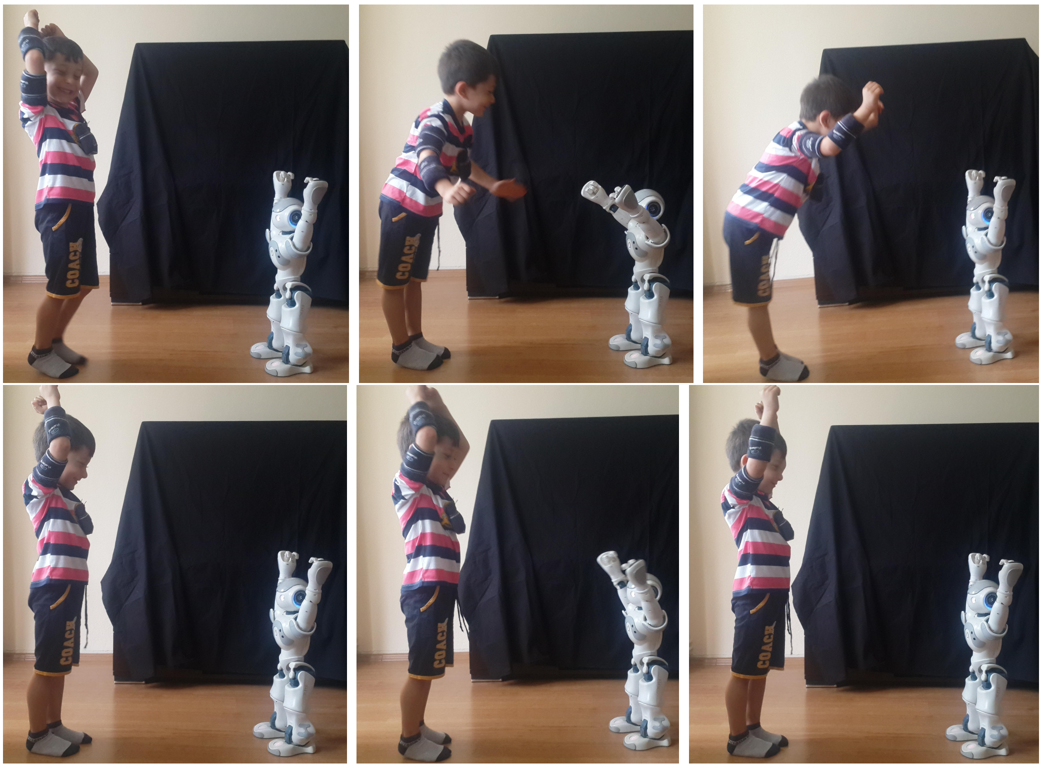

Figure 6.5. Corrected vertical extension motion after several *raise arm higher* correction feedback example 1

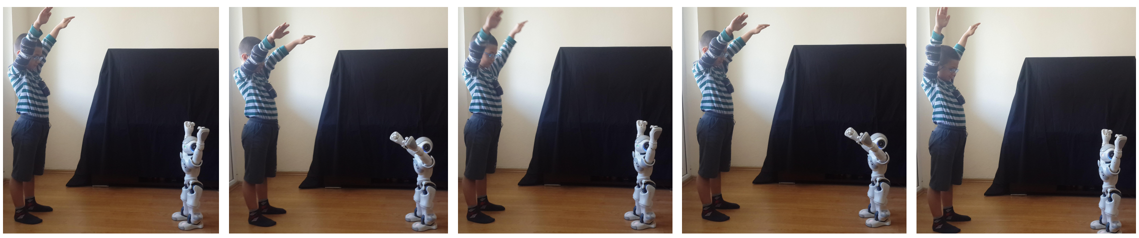

Figure 6.6. Corrected vertical extension motion after several *raise arm higher* correction feedback example 2

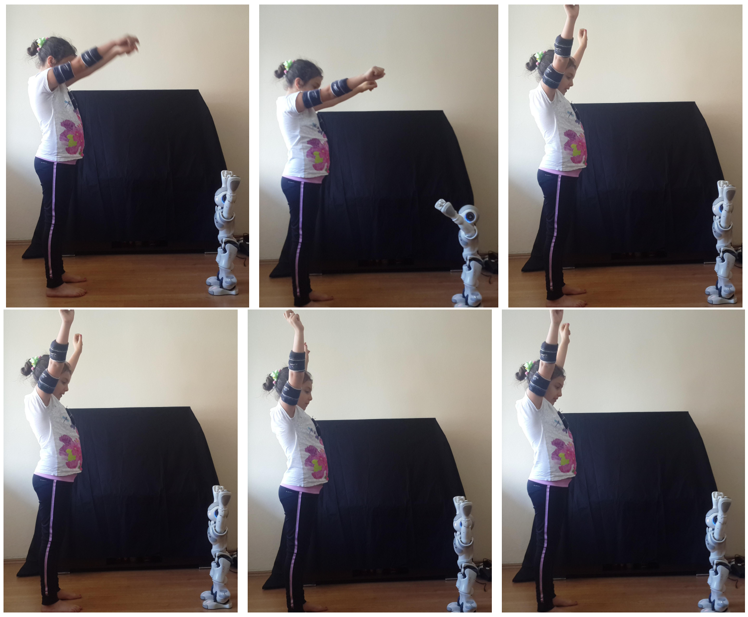

Figure 6.7. No correction during the first interaction but easily corrected in second interaction

Figure 6.8. Corrected vertical extension motion with *make elbow straight* feedback

Figure 6.9. Multiple vertical extension motion performances classified as incorrect because of higher repetitions

the robot moves its hip to the opposite side. This caused most of the children to misunderstand the motion.

(viii) Since the stretching is a hard exercise to understand most of the children raised their arms and tilted their body to the left but they did not stretch their arm together with their body. This results in high number of *raise arm higher* correction feedback because of the incorrect angle between the arm and the body.

(ix) In order to make stretching more understandable, the motion should be divided into parts. For instance, rising right arm and giving feedback then putting arm close to the right of the head then giving feedback and finally by keeping distance between arm and the face, tilting to the left and giving feedback. Because of these problems, the system most of the time partially corrected the stretching performance of the children as in Figure 6.10 and 6.11.

Figure 6.10. Partially corrected stretching example 1

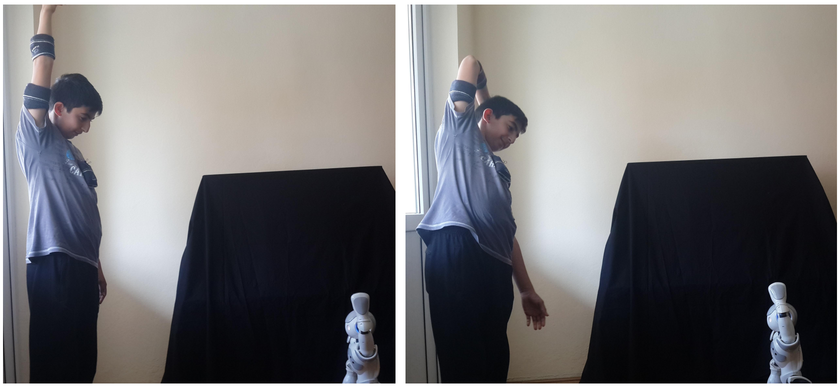

Figure 6.11. Partially corrected stretching example 2

(x) Easier motions can also be misunderstood by smaller ages as in Figure 6.12.

Figure 6.12. Performance of 4 years old child

6.1.2. Natural Gesture Imitation

Nao starts the motion flow with greetings and ends it with wiping forehead and waving hand. These natural gestures are also imitated by children without any instructions. Some examples can be seen in Figure 6.13.

6.1.3. Dance Performances

Arm dancing part is included to make the system more enjoyable for children after preliminary tests. Dance is imitated by all children except the 4 years old one

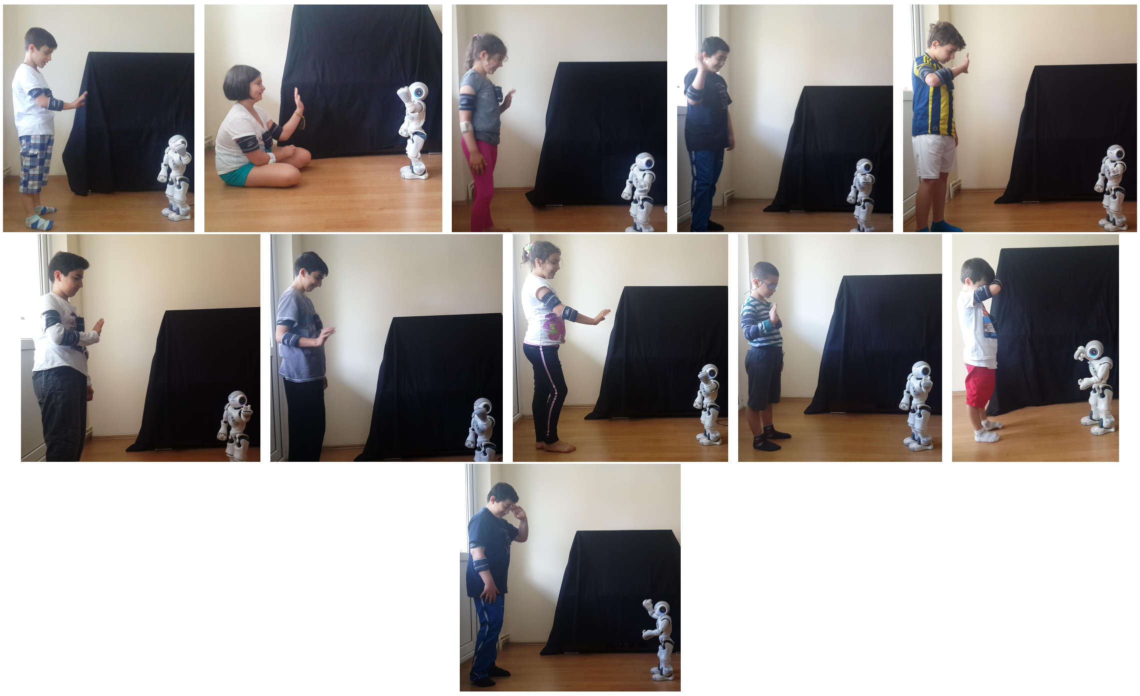

Figure 6.13. Natural gesture imitations

which stopped imitation after vertical extension performance. Even though they were sometimes observed to be shy, all of them completed at least one cycle of arm twist. Example performances can be seen in Figure 6.14.

6.1.4. False Negative Feedback Analysis

The false negative feedback ratio of the system is 11% as can be seen in Table 6.1. Wrong feedback types and their numbers are given in Table 6.4. Possible reasons behind each wrong feedback are as follows:

- "Do not tilt back/front" feedback phrases are observed to be generated on some occasions because of the involuntary pose change of the body sensor put on the chest of the child by attaching to the their t-shirts which can move while they are rising their arms as in Figure 6.15.
- "Make elbow straight" feedback is calculated using the angle between lower arm sensor and upper arm sensor. During experiments we observed that according to

Figure 6.14. Dance performances

Table 6.4. Wrong feedback and how many times it was given

Horizontal Abduction	Do not tilt front	1
	Raise arm higher	2
	Make elbow straight	7
Vertical Extension	Make elbow straight	2
	Do not tilt front	3
	Do not tilt back	3
Multiple Vertical Extension	Do not stay still, let's move	2
	Raise arm straight and two times	6
Stretching	Raise arm higher	5

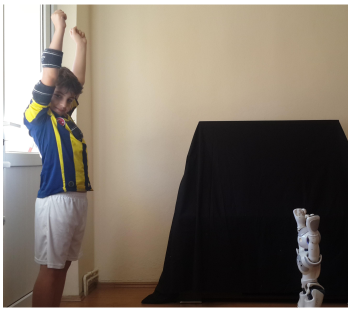

Figure 6.15. Position change of body sensor on t-shirt causes wrong *do not tilt back* feedback

the physiology of the arm of a child and flexibility of elbow joint this angle can vary. Putting sensors perfectly aligned is also a hard issue to solve. Figure 6.16 shows these bad lower sensor positioning examples.

Figure 6.16. Lower arm sensor positioning caused wrong *make elbow straight* feedback

- In multiple vertical extension motions, our system classified two motions as no motion and gave "Do not stay still, let's move" feedback when the child performed the motion with very low range of motion. This is an expected result since it is close to the no motion class. However, this observation shows us there should be one more class for two vertical extension for when it is performed with a very low range of motion.
- In multiple vertical extension motions, our system classified performances including two vertical extension with very low range of motion as wrong motion and gave "Raise arm straight and two times" feedback. This type of motions should have another class and the feedback should mention rising arms above the head instead of giving feedback on straightness and the number of motions done which

are correct in this type of incomplete performances.

- In stretching, because of the involuntary pose change of the body sensor even tough the arm was raised enough, incorrect feedback was given. In Figure 6.17 the first performance is corrected via a true feedback and then wrong feedback is given to the good performed stretch motion because of this reason.

Figure 6.17. First stretching performance corrected by an appropriate feedback but wrong feedback was given to the second performance because of involuntary pose change of the body sensor

6.2. Evaluation of System by Children

Children evaluated the interaction and the robot via several questions. Interaction is evalued in terms of enjoyableness by using seven different adjectives. The robot exercise coach is evaluated in terms of social attraction, social presence, and companionship.

6.2.1. Evaluation of the Interaction

The enjoyableness of the interaction is evaluated by using the following adjectives: “enjoyable”, “interesting”, “fun”, “satisfying”, “entertaining”, “boring” and “exciting”. Participants were asked to rate how well each adjective described the interaction on a 3-point Lakert scale, 1 describes No/Not at all or Never and 3 describes Yes/Very much/Always. Children rated the interaction highly in terms of enjoyableness as seen in Table 6.5. Only “exciting” was rated differently than the rest of the adjectives. The children who rated “exciting” with medium score commented that they saw Nao or another robot before. The usefulness of the interaction is evaluated by the question in Table 6.6 and every children found the motions very healthy and valuable for their well-being.

Table 6.5. Enjoyableness of the interaction

answers	enjoyable	interesting	fun	satisfying	entertaining	boring	exciting
yes	17	17	16	15	18		12
medium	1	1	2	3			6
no						18	

Table 6.6. Usefulness of the interaction

	Yes	No
During the exercise,were actions of the Nao useful and valuable?	18	0

6.2.2. Evaluation of the Robot

The robot’s personable qualities, such as expressing praise, continuity between sessions and politeness are best characterized by the perceived helpfulness and companionship of the robot [22]. For this reason we evaluated the perceived helpfulness and companionship of the robot via several questions explained in the following sections.

6.2.2.1. Evaluation of the Companionship of the Robot. In order to measure perceived helpfulness and the companionship of the robot we asked the children to select 5 adjectives or their opposites. Almost all of them selected the positive adjectives (see Figure 6.7).

Table 6.7. The companionship of the robot

bad	0	not loving	0	not friendly	0	not cuddly	0	cold	1
good	18	loving	18	friendly	18	cuddly	18	warm	17

6.2.2.2. Evaluation of the Robot as An Exercise Coach. The robot is evalued in terms of being a good exercise coach via five questions that can be seen in Table 6.8. Two answers to the question *Do you recommend Nao to your friends?* were observed to be negative. The children who gave this answer said they would not recommend since they do not want to share the robot with others. Afterwards, their preference of doing exercise in school or at home with or without the robot is asked and Table 6.9 shows that all of them gave a positive answer.

Table 6.8. Evaluations of the robot as an exercise coach

Question	Yes	Not much	No
I think Nao is a good exercise coach.	17	1	0
I think Nao is a good motivator of exercise.	18	0	0
Do you recommend Nao to your friends?	16	0	2
Would you like to exercise with Nao in the future?	18	0	0
Have you been motivated to exercise with Nao?	18	0	0

Table 6.9. Exercise with or without the robot

I would prefer to do exercise	with robot	18
	without robot	0

6.2.2.3. Evaluation of the Robot's Social Attributes. In order to evaluate the robot's social attributes, we measured both the social attraction towards the robot and the social presence of the robot. Social attraction questions are the modified version of the Interpersonal Attraction Scale of McCroskey et al. [42] which were also used in the study of Fasola *et. al.* [22]. Participants reported their level of agreement with the three statements given in Table 6.10.

Table 6.10. The social attraction

Question	Yes	Maybe	No
I think Nao could be a friend of mine.	18	0	0
I think I could spend a good time with Nao.	18	0	0
I would like to spend more time with Nao	17	1	0

In order to understand the perceived social presence of the robot by children two questions were asked related to machine-toy-human likeness. Results show that in machine-human comparison most of the children categorize the robot as human while in toy-human comparison they created new terms like small-human and one third of them categorized the robot as a toy (see Table 6.11 and 6.12).

Table 6.11. Machine human comparison

	Machine	Human	Both
Is Nao machine-like or human-like?	4	13	1

Table 6.12. Toy human comparison

	Toy	Human	Small Human	Both
Is Nao toy-like or human-like?	6	5	6	1

7. CONCLUSION AND FUTURE WORK

In this thesis, we presented the system design, methodology, and implementation details of a socially assistive child-robot interaction system that aims to motivate children to engage in physical exercise.

In preliminary studies we focused on the evaluation procedure in the context of rehabilitation exercises. We proposed an automated evaluation system that attempts to grade the motion performance of the child by mimicking the decision critearia of physiotherapists. Initial experiments were carried out with healthy children using a Kinect camera as monitoring tool and Nao as the robot. Evaluation of the system was done by comparing the ratings of the system with 5 physiotherapists' ratings by using Intraclass Correlation Coefficient (ICC). During evaluations and observations in the theraphy centers a limitation of the Kinect was observed (the therapist stays in front of the child during the session in order to give some objects for functional exercises which affects Kinect's vision). For this reason we switched to wearable IMUs and tested the data collection system in a rehabilitation center and made offline processing.

With our experience and observations gathered during these studies, we then began designing a real-time SAR system that will potentially better catch the child's attention and motivate him/her more to imitate the robot by serving real-time feedback as well as doing real-time evaluation of the motions of the child as opposed to offline evaluation. This new phase was conducted within an exercise coaching scenario to extend our system beyond the physical therapy domain.

Our proposed online SAR system was designed with the guidance of an exercise coach and feedback from several preliminary studies conducted with children. In order to evaluate the effectiveness of the final design an experimental study was conducted with 19 children. System performance was evaluated via accuracy of feedback that was measured to be 89%. Evaluation of the interaction and the robot were done by a questionnaire asked to the each child after their interaction with the robot. Children evaluated the interaction in terms of enjoyableness and usefulness; the robot was also rated as satisfactory in terms of exercise coaching performance, helpfullness and companionship. The social attributes of the robot were also evaluated by asking questions related to machine-human-toy likeness of the robot and most of the children categorized the robot as human-like or small-human like.

During our studies both Kinect and IMU's were used as monitoring tools. We observed certain limitations to each system: Kinect cannot monitor small motor actions like supination and the child has to stand directly in front of the camera and another person should not interfere the vision by standing in front of the child. On the other hand, IMU's were attached to the clothes which was observed to cause involuntary movements for the sensors. Another issue is that the sensor attached to the body may generate wrong angle values because of the physiology of the body. In order to generate a better recognition within SAR systems, different or combined monitoring tools can be used for different motion types and for different limbs.

We also observed a limitation of Nao in the stretching motion. Since Nao cannot move its chest without moving its hip, stretching motion was misperceived by some children. In future studies, in order to compansate for this confusion, the verbal feedback design of the robot should include precise feedback for harder motions, and these motions should be divided into some parts that will be performed and evaluated step by step.

As future work, in order to increase the activity evaluation performance of the system, the recall measure and the proposed similarity penalty algorithm can be integrated to the evaluation part together with the already used DTW distance measures. Emotion recognition can also be integrated to the system in order to understand the mood of the child and adapting feedback accordingly.

Furthermore, in order to personalize the interaction more, the robot may use the user's name at the first greeting and also during farewell. It can also provide direct feedback specific to the individual user's overall performance level and performance history during the previous interactions or trials. This mutual knowledge will potentially increase the effectiveness of the interaction.

REFERENCES

1. Feil-Seifer, D. and M. J. Mataric, “Defining socially assistive robotics”, *Rehabilitation Robotics, 2005. ICORR 2005. 9th International Conference on*, pp. 465–468, IEEE, 2005.

2. Dimitrova, M. and H. Wagatsuma, “Designing Humanoid Robots with Novel Roles and Social Abilities”, *Lovotics*, Vol. 2015, 2016.

3. Landreth, G. L., *Play therapy: The art of the relationship*, CRC Press, 2012.

4. Belpaeme, T., P. E. Baxter, R. Read, R. Wood, H. Cuayáhuitl, B. Kiefer, S. Racioppa, I. Kruijff-Korbayová, G. Athanasopoulos, V. Enescu *et al.*, “Multimodal child-robot interaction: Building social bonds”, *Journal of Human-Robot Interaction*, Vol. 1, No. 2, pp. 33–53, 2012.

5. Dahl, T. S. and M. N. K. Boulos, “Robots in Health and Social Care: A Complementary Technology to Home Care and Telehealthcare?”, *Robotics*, Vol. 3, No. 1, pp. 1–21, 2013.

6. Choe, Y.-k., H.-T. Jung, J. Baird and R. A. Grupen, “Multidisciplinary stroke rehabilitation delivered by a humanoid robot: Interaction between speech and physical therapies”, *Aphasiology*, Vol. 27, No. 3, pp. 252–270, 2013.

7. Fridin, M. and M. Belokopytov, “Robotics agent coacher for CP motor function (RAC CP Fun)”, *Robotica*, Vol. 32, No. 08, pp. 1265–1279, 2014.

8. Brooks, D. A. and A. M. Howard, “Quantifying upper-arm rehabilitation metrics for children through interaction with a humanoid robot”, *Applied Bionics and Biomechanics*, Vol. 9, No. 2, pp. 157–172, 2012.

9. Strohrmann, C., R. Labruyère, C. N. Gerber, H. J. van Hedel, B. Arnrich and

G. Tröster, "Monitoring motor capacity changes of children during rehabilitation using body-worn sensors", *Journal of neuroengineering and rehabilitation*, Vol. 10, No. 1, p. 83, 2013.

10. Zannatha, J. M. I., A. J. M. Tamayo, Á. D. G. Sánchez, J. E. L. Delgado, L. E. R. Cheu and W. A. S. Arévalo, "Development of a system based on 3D vision, interactive virtual environments, ergonometric signals and a humanoid for stroke rehabilitation", *Computer methods and programs in biomedicine*, Vol. 112, No. 2, pp. 239–249, 2013.

11. Ranatunga, I., M. Beltran, N. A. Torres, N. Bugnariu, R. M. Patterson, C. Garver and D. O. Popa, "Human-Robot Upper Body Gesture Imitation Analysis for Autism Spectrum Disorders", *Social Robotics*, pp. 218–228, Springer, 2013.

12. Greczek, J., E. Kaszubski, A. Atrash and M. Matarić, "Graded cueing feedback in robot-mediated imitation practice for children with autism spectrum disorders", *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*, pp. 561–566, IEEE, 2014.

13. McCarthy, C., J. Butchart, M. George, D. Kerr, H. Kingsley, A. M. Scheinberg and L. Sterling, "Robots in Rehab: Towards socially assistive robots for paediatric rehabilitation", *Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction*, pp. 39–43, ACM, 2015.

14. Tsiakas, K., "Facilitating Safe Adaptation of Interactive Agents using Interactive Reinforcement Learning", *Companion Publication of the 21st International Conference on Intelligent User Interfaces*, pp. 106–109, ACM, 2016.

15. Calderita, L. V., P. Bustos, C. S. Mejías, F. Fernández and A. Bandera, "Therapist: towards an autonomous socially interactive robot for motor and neurorehabilitation therapies for children", *Proceedings of the 7th International Conference on Pervasive Computing Technologies for Healthcare*, pp. 374–377, ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering),

2013.

16. Martín, A., J. C. González, J. C. Pulido, Á. García-Olaya, F. Fernández and C. Suárez, "Therapy Monitoring and Patient Evaluation with Social Robots", *Proceedings of the 3rd 2015 Workshop on ICTs for improving Patients Rehabilitation Research Techniques*, pp. 152–155, ACM, 2015.

17. Ramgoolam, D., E. Russell and A. B. Williams, "Towards a social and mobile humanoid exercise coach", *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*, pp. 274–275, ACM, 2014.

18. Matsusaka, Y., H. Fujii, T. Okano and I. Hara, "Health exercise demonstration robot TAIZO and effects of using voice command in robot-human collaborative demonstration", *RO-MAN 2009-The 18th IEEE International Symposium on Robot and Human Interactive Communication*, pp. 472–477, IEEE, 2009.

19. Park, C., J. Kim and J.-H. Kang, "Robot Social Skills for Enhancing Social Interaction in Physical Training", *The Eleventh ACM/IEEE International Conference on Human Robot Interation*, pp. 493–494, IEEE Press, 2016.

20. Kidd, C. D. and C. Breazeal, "A robotic weight loss coach", *Proceedings of the national conference on artificial intelligence*, Vol. 22, p. 1985, Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, 2007.

21. Kidd, C. D. and C. Breazeal, "Robots at home: Understanding long-term human-robot interaction", *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*, pp. 3230–3235, IEEE, 2008.

22. Fasola, J. and M. Mataric, "A socially assistive robot exercise coach for the elderly", *Journal of Human-Robot Interaction*, Vol. 2, No. 2, pp. 3–32, 2013.

23. Howard, A. M., L. Roberts, S. Garcia and R. Quarells, "Using mixed reality to map human exercise demonstrations to a robot exercise coach", *Mixed and Augmented*

Reality (ISMAR), 2012 IEEE International Symposium on, pp. 291–292, IEEE, 2012.

24. Görer, B., A. A. Salah and H. L. Akın, "A robotic fitness coach for the elderly", *Ambient Intelligence*, pp. 124–139, Springer, 2013.

25. Ros, R., I. Baroni and Y. Demiris, "Adaptive human–robot interaction in sensorimotor task instruction: From human to robot dance tutors", *Robotics and Autonomous Systems*, Vol. 62, No. 6, pp. 707–720, 2014.

26. Wilk, R. and M. J. Johnson, "Usability feedback of patients and therapists on a conceptual mobile service robot for inpatient and home-based stroke rehabilitation", *Biomedical Robotics and Biomechatronics (2014 5th IEEE RAS & EMBS International Conference on*, pp. 438–443, IEEE, 2014.

27. Wainer, J., D. J. Feil-Seifer, D. A. Shell and M. J. Mataric, "Embodiment and human-robot interaction: A task-based perspective", *Robot and Human interactive Communication, 2007. RO-MAN 2007. The 16th IEEE International Symposium on*, pp. 872–877, IEEE, 2007.

28. Guneysu, A., R. D. Siyli and A. A. Salah, "Auto-evaluation of motion imitation in a child-robot imitation game for upper arm rehabilitation", *Robot and Human Interactive Communication, 2014 RO-MAN: The 23rd IEEE International Symposium on*, pp. 199–204, IEEE, 2014.

29. Csikszentmihalyi, M., *Beyond boredom and anxiety.*, Jossey-Bass, 2000.

30. "EXL-S3 Inertial Measurement Units", http://www.exelmicroel.com/index.html, accessed at June 2016.

31. Guneysu, A., B. Arnrich and C. Ersoy, "Children's rehabilitation with humanoid robots and wearable inertial measurement units", *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2015 9th International Conference on*, pp.

249–252, IEEE, 2015.

32. "Dynamic Time Warping", https://en.wikipedia.org/wiki/Dynamic_time_warping, accessed at June 2016.

33. Norkin, C. C. and D. J. White, *Measurement of joint motion: a guide to goniometry*, FA Davis, 2009.

34. Pot, E., J. Monceaux, R. Gelin and B. Maisonnier, "Choregraphe: a graphical tool for humanoid robot programming", *Robot and Human Interactive Communication, 2009. RO-MAN 2009. The 18th IEEE International Symposium on*, pp. 46–51, IEEE, 2009.

35. Müller, M., "Dynamic Time Warping", *Information Retrieval for Music and Motion*, pp. 69–84, Springer Berlin Heidelberg.

36. "Intraclass Correlation and Related Methods", http://www.john-uebersax.com/stat/icc.htm, accessed at July 2013.

37. Bulling, A., U. Blanke and B. Schiele, "A tutorial on human activity recognition using body-worn inertial sensors", *ACM Computing Surveys (CSUR)*, Vol. 46, No. 3, p. 33, 2014.

38. "The Instrument Control Toolbox Bluetooth Interface of Matlab", http://www.mathworks.com/help/instrument/bluetooth.html, accessed at June 2016.

39. Preece, S. J., J. Y. Goulermas, L. P. Kenney, D. Howard, K. Meijer and R. Crompton, "Activity identification using body-mounted sensors—a review of classification techniques", *Physiological measurement*, Vol. 30, No. 4, p. R1, 2009.

40. "The NAO-Base Robot Behaviors", http://funlab.nd.edu/the-nao-base/, accessed at June 2016.

41. Bickmore, T., D. Schulman and L. Yin, "Maintaining engagement in long-term interventions with relational agents", *Applied Artificial Intelligence*, Vol. 24, No. 6, pp. 648–666, 2010.

42. McCroskey, J. C. and T. A. McCain, *The measurement of interpersonal attraction*, Taylor & Francis, 1974.