NAIST-IS-DD1061004

Doctoral Dissertation

# A Study of Assist-As-Needed Robotic Training Based on Reinforcement Learning

Chihiro Obayashi

March 13, 2014

Department of Bioinformatics and Genomics
Graduate School of Information Science
Nara Institute of Science and Technology

A Doctoral Dissertation
submitted to Graduate School of Information Science,
Nara Institute of Science and Technology
in partial fulfillment of the requirements for the degree of
Doctor of ENGINEERING

Chihiro Obayashi

Thesis Committee:

| | |
|---|---|
| Professor Kazushi Ikeda | (Supervisor) |
| Professor Tsukasa Ogasawara | (Co-supervisor) |
| Visiting Professor Tomohiro Shibata | (Co-supervisor) |

# A Study of Assist-As-Needed Robotic Training Based on Reinforcement Learning*

Chihiro Obayashi

**Abstract**

This study proposes a novel robotic trainer for motor learning. It is user-adaptive inspired by assist-asneeded principle well known in the filed of physical therapy. Most of the previous studies in the field of robotic assistance of motor skill learning requires predetermined desired trajectories which was not examined intensively if they were optimal for each user. Furthermore, it has been known as guidance hypothesis that humans tend to rely too much on external assist, resulting in interference with internal feedback necessary for motor skill learning. A few studies proposed such a system that adjust its assist-strength according to the user's performance in order to prevent the user from being too much relying on the robotic assistance. There are, however, problems in those studies; The physical model of the user's motor system is required, which is inherently difficult.

In this study, I propose a framework for such a robotic trainer that is user-adaptive, and that does not require a specific desired trajectory nor the physical model of the user's motor system, which is achieved by a model-free reinforcement learning. We chose darts throwing as an example motor-control task as it is one of the simplest throwing task, and its performance can be easily and quantatively measured by score. Training experiments with novices, aiming at maximizing the score of the darts and minimizing the physical robotic support demonstrate the feasibility and the plausibility of the proposed framework.

*Doctoral Dissertation, Department of Bioinformatics and Genomics, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-DD1061004, March 13, 2014.

**Keywords:**

Assistive Robotics, Motor Skill Learning, Human-Robot Interaction, Throwing motion, Reinforcement Learning

# 強化学習を用いたユーザ適応型運動スキル学習支援ロボットに関する一研究*

大林 千尋

**内容梗概**

近年，ユーザの動作データに基づいたパワーアシストや視聴覚フィードバック等の運動学習支援システムの研究が盛んに行われている．しかし，多くの支援システムは予め用意された運動軌道の学習を目的としており，ユーザの多様性を許容しにくい．本来ヒトは支援に依存し易く，ユーザの内在的なフィードバックを阻害し運動スキル学習が妨げられる．学習状態を推定し支援量を調整する支援システムが提案されているが，あらかじめユーザの身体モデルが必要であり，ヒトの運動制御系全体を精密にモデリングすることは困難である，という課題があった．

本論文では，運動スキルなどの運動学習指標を報酬関数に組込んだモデルフリー型の強化学習により，非軌道追従型支援をユーザ適応的に行うロボット支援フレームワークを提案し，最も単純な投擲運動の一つであるダーツ投げ運動を題材に，その有効性を示す．まず，支援方法の設計のために熟達者と非熟達者の動作データや表面筋電位による比較を行い，投げ動作中のユーザの肩や肘の移動量が，ダーツ投げ熟達度の定量指標となり得ることを確認する．次にこの指標を状態変数とし，非軌道追従型のロボット支援法を方策とし，ダーツのスコア最大化と支援量最小化を報酬関数としたモデルフリー強化学習システムを設計する．行動実験を実施し，提案手法がダーツ投げの運動学習に最も有効であることを示した．

**キーワード**

Assistive Robotics, Motor Skill Learning, Human-Robot Interaction, Throwing motion, Reinforcement Learning

---

*奈良先端科学技術大学院大学 情報科学研究科 情報生命科学専攻 博士論文，NAIST-IS-DD1061004, 2014 年 3 月 13 日．

# Contents

# List of Figures

# List of Tables

# 1. Introduction

In future, robots are expected to be closely involved with human life, rather than just being a simple machine. We consider that robots need to perform assistive tasks such as playing the role of an instructor in the case of motor learning. An adaptive assist robot that uses physical interaction to help human motor learning is expected to be developed. In a scenario of motor learning assisted by an instructor, the instructor's assist method can be adapted to the performance and physical characteristics of the learner. For example, imagine a robot assisting a learner to ride a bicycle by holding the carrier of the bicycle. Here, the robot needs to adjust the assistance to suit the learner's characteristics: physical characteristics, training experiences, etc.

Robots developed in previous studies for assisting motor learning are based on the idea that the best way to learn a movement is to track the underlying movement for a motor task [1], [2], [3], [4], [5]. Although the learner's physical characteristics and motor learning experience was different, the learner automatically tracked simple uniform trajectories provided by a special robot and learned these trajectories. Developing an assistive robotic system that can adjust to individual differences in the learner's physical and training experience factors is a challenging research topic. The basic technology challenges were related to motion sensing, motion analysis, and development of the control system of the assist robots. Another challenge involved in developing such a system is related to human motor control and motor learning. Therefore, this topic is being researched from the perspective of not only robotics and biomechanics but also neuroscience and rehabilitation/sports science.

In rehabilitation, researchers have focused on re-learning of motor functions by brain injury patients. They have developed a robot-assisted system based on quantification of training. This research field is called rehabilitation robotics or neuro-rehabilitation. In rehabilitation robotics, such systems have been studied separately from the perspective of developing movement-assisting hardware for physical assistance and applications using such hardware. Different motor learning assist hardware systems have been proposed based on a tracking method [1], myoelectric potential information control [2] and adaptive learning assist that tracks patient performance [6]. Many rehabilitation robotics studies have focused on cases involving brain damage to take advantage of the plasticity of the brain [7]. By training, a limb that has lost its function can be forced to generate sensory stimulation activity in the brain; owing to the plasticity of the brain, the somatotopic representation (map) of movement changes and somatosensory potentials in the normal part of the brain are intercepted so that the limb function is recognized [8], [9]. In recent years, robot-assisted training utilizing the plasticity of the cerebellum has also been proposed [10]. By adding a motion stimulus, a patient can be forced to use and learn the internal model of a new motor task [11].

The high adaptation ability of humans to a movement environment could be problematic when learning the internal model of a new motor task. For example, consider a person using a tool as though it is a physical part of the body and is known to realize movement to achieve the purpose of movement [12], [13]. One of the motor learning properties called the “after effect” was demonstrated. Consider a scene in which we attempt to learn the linear motion of the upper extremity in a certain force field. When the force field is removed suddenly, the upper extremity’s trajectory exhibits curves. This is called the “after effect” [14]. This is because when the force field is changed suddenly, the internal model of the next force field is not reshuffled. Therefore, it is thought that the adaptation to a training environment with assistance obstructs the realization of smooth transfer of the training in a real environment.

In sports science, the phenomenon of motor learning is known as guidance hypothesis [15]. When novices practice with too much assistance in the early stage of motor learning, they tend to rely on the assistance and are not able to give their best performance without the assistance. Therefore, it is desirable to develop a computer agent that automatically adapts the assistance to the trainer based on quantitative indices of skillful movements that clarify the difference between experts and novices.

The adaptability of robotic assistance for motor skill learning was investigated with the aim of applying it to robotic assistance for humans with/without stroke. Most previous studies on robotic assistance have required predetermined desired trajectories; it has not been examined intensively whether these trajectories were optimal for each person. Furthermore, it has been inferred from guidance hypothesis that humans tend to rely excessively on external assistance, resulting in interference with internal feedback that is necessary for motor skill learning. A few studies proposed a system that adjusts its assist-strength according to a person's performance in order to prevent the person from relying excessively on robotic assistance. However, these studies have some problems; the physical model of the person's motor system is required, which is inherently difficult.

## 1.1 Problem of Motor Skill Learning

Under the new movement environment, motor learning is explained by the knowledge of computational neuroscience to build the internal model of the training environment dynamics, including the skeletal system of the body dynamics. When we assist the learning of the internal model for a new motion, it is necessary to consider the following problems: difference in the motor learning experience, physical differences (individual differences), and motor learning properties. Regarding individual differences, even for persons of the same age and gender, the body characteristics are different depending on everyday training. Actually, in the study, the average muscle mass is calculated by linear regression using the height and weight because one part of the body (hand, leg, etc.) had larger variance. Therefore, it is difficult to estimate the quantity of muscle mass for each person. In addition, the learning speed is also different in individual persons [16], [17].

Regarding motor learning properties, we can consider the effect called "after effects" as one of the properties of motor learning [14]. When the disturbing force field is unexpectedly removed, subjects make erroneous movements in directions opposite to the perturbing forces. The force field is changed suddenly, and a learner is explained when reshuffing is not carried out well for the internal model of the next force field. In addition, the person uses a tool as though it is a physical part of the body. This embodiment function is known to realize movement to achieve an action [12], [13]. Because such a property exists, it is necessary to prevent it from being used in some learning environment as the assistance tool realizes an action.

Thus, it is necessary to consider the properties of motor learning (individual differences, and so on) when we design the assist method or machine for motor learning.

## 1.2 Assist Framework

The robotic assistance system for motor learning started with the development of an exoskeleton robot system assisting a person with a high-risk training task that involves a physical burden that is too heavy for a person to lift, including the lifting of heavy weights. The most pioneer work in this field is Hardiman (Human Augmentation Research and Development Investigation), which was developed by U.S. General Electronics in 1965. This exoskeleton is an external skeleton-type robot aimed at lifting heavy things [18],[19]. However, it was not commercialized because it had low operability. Recent wearing-type exoskeletons for improving the power of the person are compact and have high response because the basic technology used has advanced, the control unit and the actuator have been downsized, and the operability has been improved [20], [21], [22]. A rehabilitation robotic system for lower limbs was developed by MIT Media lab [23]. A wearable walking assistive robotic system (called HULC) was developed by Sarcos Company for military use [24], [23]. In Japan, the most representative study was the robot suit HAL, which is versatile to improve a non-healthy person to a healthy person; it was developed by Sankai Lab in Tsukuba University [22],[25]. In addition, a power assist suit for older farm workers was developed by Endo lab in Tokyo University of Agriculture and Technology [26]. WAS-LiBERo, which assists the waist joint at the time of lifting a farm product collection gauge and the hip joint at the time of walking with the gauge, was developed by Yagi lab in Wakayama University [27], [28], [29]. A muscular strength assistance harness using an elastic body for the waist joint was applied to firefighting work and skill assistance; it was developed by Tanaka team in Hokkaido University [30]. An exoskeleton that could assist the main joint group of the whole body using a pneumatic actuator to realize safe, natural care movement such as the lifting movement of the care-giver was developed by Yamamoto lab in Kanagawa Institute of Technology [31], [32]. An assist suit for the hip joint, which used CPG for a control system for the purpose of a natural walk assist, was developed by Hashimoto lab in Shinsyu University [33].

In recent years, car companies have entered in the assistive robotic system research field. HONDA Motor Co., Ltd., announced that they are conducting research on rhythm walk assistance and its development to downsize the sensor control system and hence realize a natural walk. In 2008, they proposed to improve the mobility of a person [21]. Honda collaborated with the National Center for Geriatrics and Gerontology in 2012. Honda proposed the care preventive effect of walking assistance using the assist suit. In 2011, TOYOTA Motor Co., Ltd., announced the development of many assist robots for walking assistance to rehabilitate person [20]. Toyota has conducted a clinical trial since the development of assist robots in collaboration with Fujita Health University.

In this section, various assist robots have been introduced. Most studies are based on hardware development. There are a few studies based on the control problem to operate the assist suits or the robots. In the case of the control of the assist suit for motor learning, it is necessary to determine the function for the estimate of "how much quantity of assistance is necessary." The problem to estimate the intention of movement has been discussed in the field of the development of a high-performance artificial arm and wheelchair controller by using biological information. There are a few studies on the estimation of the intention of the person by using movement information [34], [35], [36]. There are also a few robot control studies based on electroencephalographic informationThere are also a few robot control studies based on electroencephalographic information [37], by using surface electromyography information [38], [39], [40], [41], [42]. However, in recent years, the function for the estimate of "how much quantity of assistance is necessary" is just beginning to be discussed in the field of rehabilitation robotics; this will be the key technology for more useful assistive robotic systems.

In initial rehabilitation robotics, the purpose of the research and the development were aimed at reducing the physical load of a therapist's daily work by robotic automation. Specifically, these developed systems aimed at robotic automation for training therapy carried out by a therapist on a non-healthy person.

Next, an assistive robotic system is not used to perform power assist for the therapist, such as PAS. The rehabilitation robotic system was developed to assist motor learning for non-healthy persons. The system was of the deferment type for safety purposes. In addition, a robot to realize tracking-based assistance for motor learning restrains the body of persons. The system traces the prepared trajectory for a non-healthy person to realize the task-oriented assistance, which is a famous manual rehabilitation method. Therefore, motor learning assistance with the robotic system for exclusive use of the movement was performed. However, this method is not versatile to be used for other movements.

Concrete studies on the rehabilitation robot system are as follows. The MIT Manus system was proposed for assisting point-to-point upper limb movement on the horizontal plane [1]. MIME was proposed to be able to perform both upper limb mirror training on the hand space by a 6-DOF manipulator [2]. Arm Guide was proposed for assisting the reaching movement of arm space. Arm Guide has a 3-DOF flexibility arm comprised of a straight slide-type actuator [3]. Next, a motor learning assistive robotic system for locomotion was also developed. The assist robot performs power assist for non-healthy persons to realize the joint trajectories for locomotion [4], [5]. It was shown that there was a constant effect by carrying out training early in the stroke, and these systems performed the role of indicating quantitative scientific grounds for neuro-rehabilitation [43]. However, in the study on the walk assist robot, in comparison with the therapy by the therapist using Body Weight Assisted Treadmill Training, the problem was that the effect around the training time was low [44].

In contrast, it was thought that the problem is uniform assist to a different non-healthy person having individual skeletal systems for lower limbs. The training system, which provided desired trajectories in consideration of the physical difference of the learner, was suggested [45]. The assistive robotic system using elastic devices like springs and the assistive robotic system having virtual impedance control by software were developed to realize the play for individuality of the learner [6], [46], [47]. Only animal experiments suggested that tracing the trajectory with play by the assist robot is effective [48]. However, even these methods were not able to suggest more possibilities than the effects of the therapy by the therapist using Body Weight Assisted Treadmill Training in humans. Recently, in order to solve this problem, a study that aimed at training effect improvement by offering more widespread assisting circumstances was conducted.

This robotic assistance problem is that the robot assistance performs repeated monotonic movements. Therefore, the learner's trial and error learning and opportunity to challenge are easily spoiled. It is thought that the training of the learner will decrease because of such factors. The learner tends to rely on the robotic assist gradually. In recent years, various approaches for the rehabilitation robotic system's problem were proposed. The first is the study of the communication between learners and the robot to solve this problem. This robot learned the distance to decrease the psychological burden on each learner for fitting the learner [49], [50], [51]. The second is the study that introduced the presentation using sound, visual information, and other modalities [52], [53], [54], and the inclusion of the gaming property into the training task with robotic assistance [55], [56], [57]. Furthermore, the robot providing adaptive physical assist to the state of the learner is studied. This study is in coordination with the quantity of assistance based on the performance of the learner [6], [46], [58], [10], [59].

Krebs et al. proposed performance-based robotic assistance [6]. Their system regulated the quantity of assistance by manual adjustment. Emken et al. expressed a combination of an error and voluntary display of the tip of the foot position of the walking subjects for the cost function based on Assist-As-Needed strategy [58], [60]. They designed the learning control system for assisting the motion of " put up foot " to coordinate the assist force. They used the model of the brain-muscle-skeletal system of a human in a control system to estimate the output force of the lower limbs. They suggested that when the assist ratio was set to perform a large force by a human, the motor learning effect was improved. Their results the after effect in the adaptive condition were smaller than the constant assist condition. It was suggested that even if they were assisted by an adaptive assist robot, they can learn the internal model in the environment without robotic assistance. However, they have not shown that the adaptive robotic assistance promotes motor learning. In addition, it is necessary to identify the learner's musculoskeletal model, and there are many identification parameters to coordinate the model because rehabilitation training is needed to apply various persons quickly. The model-based approach spent time to identify the parameters before each training. Therefore, we think that the model-based control approach was not suitable to use the rehabilitation system.

In these circumstances, we propose a novel framework of the robot assistive robotic system based on the Assist-As-Needed strategy. The proposed framework adopts a model-free algorithm that does not need the model construction of each learner. Specifically, we will use an online-type reinforcement learning algorithm that can realize the Goal-Oriented assist for the learner [61], [62]. In addition, our proposed assist framework realizes the minimum contact assist for examination of the possibility of the assist with not an exclusive machine but with a general robot. We build non-trajectory-based movement assistance to promote voluntary training participation of the learner and avoid relying on the robot.

## 1.3 Organization of Dissertation

This dissertation is organized as follows. Section 2 proposes a new approach to realizing non-trajectory-based adaptive robotic support system. Section 3 describes the comparison between experts and novices in darts and the results to measure the dart trajectory. Section 4 describes the development of an adaptive learning assistive robotic system for dart -throwing and the experimental setup to investigate the feasibility and plausibility of our approach. We summarize this study with detailed discussions and recommendations for future research in section 5.

# 2. Proposed Adaptive Robotic Assist Framework

## 2.1 Specification

Providing excessive assistance may adversely affect learning; therefore, a commonly stated objective of active assist training is to provide ”assistance-as-needed” (AAN), which means to provide the participant with only as much assistance as is needed to accomplish the task (sometimes called ”faded guidance” in motor learning research). Examples of strategies to encourage participant effort and self-initiated movements include allowing some error variability around the desired movement using a dead band (an area around the trajectory in which no assistance is provided), triggering assistance only when the participant attains a force or velocity threshold, making the robot compliant, and including a forgetting factor in the robotic assistance scheme. A pioneering study was conducted by Krebs et al. [6] who used a robotic device with the robotic assist optimized for a person on the basis of the developer's hand coding. In recent studies, the developer has designed the cost function and created a mechanism for determining the assist (e.g., assist force) in order to minimize the cost function (e.g., error between desired trajectory and person with robot one) automatically [60], [46]. In such approaches, it is necessary to carefully design a model of the neuro-musculoskeletal system of a person. However, it is difficult to develop such a model precisely. In addition, the training entails high physical strain and a long setting time. In addition, most assistive robots have the ability to follow a trajectory traversed by a person. Such assistive robotic systems are called ”trajectory-based assistive robotic systems”. The advantage of such systems is that a person can learn simple motion. However, learning basic motion may not be the primary requirement of every person because people have different physical characteristics and motion skill levels. Such systems may not have an across-the-board learning phase from relearning of motor skills to their advancement. On the other hand, a non-trajectory-based hands-on approach involving minimally assistive robot training for proprioception enhancement based on the AAN paradigm is proposed[63]. This approach provides minimal assistance passively. In other words, movements are assisted, not enforced, by the robotic assist. However, this approach was not evaluated in sufficient detail, as the aforementioned study incorporated visual feedback in an experiment for evaluating the proposed assist.

## 2.2 Our Assistive framework

In this study, I propose an assistive robotic system, i.e., an AAN-based assist robotic system using a model-free online reinforcement-learning algorithm and non-trajectory assist, in order to provide assistance to persons with different needs. The key features of the framework are summarized below.

**Task-goal oriented** In general, it is not trivial to predetermine a desired trajectory for motor skill learning because each person has his/her own motor skills and control system. Since one of the most important aims of motor skill learning is to accomplish a task, the aim of the robotic trainer should be task-goal oriented, which requires a method for measuring a person's achievement (performance) on the task.

**Assist-as-needed** According to guidance hypothesis, humans tend to be over-reliant on external assistive feedback, which reduces the efficacy of the internal feedback necessary for motor skill learning. Therefore, the robotic trainer should adjust its assist-strength according to the measured task performance of a person, i.e., it should decrease its assist-strength when the person's performance increases, and vice-versa.

**Model-free** It is nontrivial to define the optimal throwing trajectory for each person in advance, owing to individual differences in body dynamics and in the neural controller. Therefore, the robotic trainer should employ a model-free assist algorithm.

**Minimum constraint** It is also nontrivial to determine the optimal assisting policy for each person in advance; the robotic trainer should attempt to minimize the constraints on a person's motion, which would enhance the safety of the system.

### 2.2.1 Task-goal oriented

In general, it is not trivial to predetermine some desired trajectory for motor skill learning because each person has his/her own motor skills and control system. Since one of the most important aims of motor skill learning is to accomplish a task, the aim of the robotic trainer should be task-goal oriented, which requires a method for measuring a person's achievement on the task. Trajectory-based rehabilitation robots were developed for the purpose of re-learning lost motor functions in the event of brain damage. The required assistance is provided by joint torque to generate a motion trajectory for each joint of the lower extremities, and to realize walking motion by tracking the desired trajectories. A typical walking pattern was presented for each person. However, the human musculoskeletal system exhibits individual variations. The purpose of walking rehabilitation is to enable subjects to walk by themselves. It is unlikely that learning to track a typical walking trajectory would serve this purpose. Therefore, assistive robots need to adjust the assist to achieve not only the learning of a motion trajectory but also the essential purpose.

### 2.2.2 Assist-as-needed

According to guidance hypothesis, humans tend to be over-reliant on external assistive feedback, which reduces the efficacy of the internal feedback necessary for motor skill learning. Therefore, the robotic trainer should adjust its assist-strength according to the measured performance task of a person, i.e., it should decrease its assist- strength when the person's performance increases, and vice-versa.

Guidance hypothesis experimentally describes how assistance inhibits learning in sports psychology [64]. Most recently proposed adaptive assistive robotic systems are based on this hypothesis. Because providing too much assistance may adversely affect learning, a commonly stated objective of active assist training is to provide "assistance-as-needed", which means to provide the participant with only as much assistance as is needed to accomplish the task (sometimes termed "faded guidance" in motor learning research). Examples of strategies to encourage participant effort and self-initiated movements include allowing some error variability around the desired movement using a dead band (an area around the trajectory in which no assistance is provided), triggering assistance only when the participant attains a force or velocity threshold, making the robot compliant, and including a forgetting factor in the robotic assistance, as reviewed below. When we perform an action under a certain environment, we learn an internal model of the environment by performing a repeat trial. This is shown as the nature of motor learning to perform reaching movements between two points by using a robot system, which is capable of generating any force field. In the experiment of Shadmehr and Mussa-Ivaldi et al.[14], when a person is repeatedly exposed to a robot-generated force field applied to the hand (forces as a function of hand position and/or hand velocity) that systematically disturbs limb motion, he/she is able to recover his/her original kinematic patterns over a short period of practice. The subject does this by cancelling the disturbance with an appropriate preplanned pattern of forces. This is a form of feed-forward control that is revealed by characteristic after-effects: when the disturbing force field is unexpectedly removed, subjects make erroneous movements in directions opposite to the perturbing forces. Adaptation and its related after-effects have been demonstrated for different types of force fields, simple position-, velocity- and acceleration-dependent force fild to Coriolis forces caused by moving in a rotating room to skew-symmetric" curl "field that produce forces in direction perpendicular to the velocity of hand. In the experiment of Osu et al. [65], when a learner learns two actions with the corresponding visual stimuli, he/she estimates the force field by using information of the visual stimulus and performs an action based on estimated its. Therefore, we change the internal model corresponding to situations and perform actions by using it. Thus, a person tends to obtain the internal model of the environment and uses it to achieve his/her objectives. This means that with continued physical assistance from the robot, a person will learn to operate only in environments where robotic assistance is available. In other words, the person become dependent on the assist. For example, in [66], the impedance control of the learning assist for walking was used, but the energy consumption of learners in training 60 percent compared with the help of a therapist was small. In [67], learning assist of musculoskeletal upper extremities is studied, but since the robot assist is performed unnecessarily, the driving force of the learner is essentially decreased compared with before training. These findings suggest what might be termed the "Slacking Hypothesis": a robotic device could potentially slow down recovery if it encourages slacking; i.e., a decrease in motor output, effort, energy consumption, and/or attention during training. If you do assist the body in motor learning, it is not possible to ignore this effect. In order learn the action of the learner himself corresponding to the real environment, the robot uses a method to match the needs of the learner assist.

For example, in rehabilitation for walking, the goal of the patient is to be able to walk without assistance. Even if the patient can walk with any assistance, the purpose of the rehabilitation cannot be achieved because the purpose of the rehabilitation is that the patient should eventually be able to move as much as possible without assist. In manual rehabilitation, by checking the status of the patient, the therapist performs assistance. This is called Assist-As-Needed by rehabilitation robot researcher. Therefore, it is necessary to tailor the assist robot to the learning of the patient. An adaptive assist robot tailored to patients was proposed. The adaptation of the control parameters based on each patient's physical characteristics and learning character to adjust the assist automatically is an essential advantage of the adaptive assist system.

### 2.2.3 Model-free

It is nontrivial to define the optimal throwing trajectory for each person in advance, owing to individual differences in body dynamics and in the neural controller. Moreover, the learners' age and gender may vary. The musculoskeletal system also varies from individual to individual. Furthermore, each person has a different movement experience. Thus, because learners have varying characteristics, it is difficult to precisely model the motor learning system of the brain based on the training experience and accurate modeling of the musculoskeletal system in order to provide rich adaptability. Further, in order to make the system ready for a training environment in a short period, it is necessary to use an easy experimental protocol with a simple experimental system. The model-free approach adapted successively to the learner in every trial by a data-driven approach. We can design the adaptive system without the need for prior information of the learner. It is not necessary to identify the parameters and modeling in advance; the model-free adaptive algorithm can perform the training of individuals in a short time. For existing rehabilitation robots, most studies built the control system based on iterative learning control (ILC) to determine the amount of assist tailored to the learner [68]. Because these assistive robotic systems were used for the target trajectory, their adaptability is realized by repeating the control input to track the target trajectory (assist). Duschau-Wicke et al. designed a control system based on iterative learning control as cooperative-assist. Their system estimated the appropriate knee joint torque auxiliary amount of learning assist by Lokomat [46]. Crespo et al. and Emken et al. obtained an estimate of the amount of assist by using a combination algorithm of musculoskeletal and ILC [69], [60]. Their method needs modeling of the contact between the robot and the human musculoskeletal system of the foot. However, it is necessary to identify the learner model by their technique beforehand, and there are many identification parameters to coordinate because the rehabilitation training needs to be applied to various persons quickly. The model-based approach needs the identification of the parameter before the training. Therefore, the robotic trainer should employ a model-free assist algorithm.

### 2.2.4 Minimum constraint

It is also nontrivial to determine the optimal assisting policy for each person in advance; the robotic trainer should attempt to minimize the constraints on a person's motion, which would also enhance the safety of the system. A human's actions are slightly different in each trial although it is desired to perform the same action. This occurs because of noise called the signal-dependent noise: SDN mixed in control command from the brain. As one possible factor, it is considered that a human performs a variety of actions to cope with uncertain environments or situations. An assistive robot can tolerate these differences in operation by imposing a weak constraint. For example, in the training of patients with a spinal cord injury, when the robot completely assists the learner, the learner's nervous system does not shown voluntarily activities. In such cases, the learning will fail despite the efforts of the robotic assistance. The learner's attempt to voluntarily recognize his/her own performance is important in motor learning. The robotic assistance is not fixed completely; it should provide assistance tailored to the learner so that the learners are able to face the challenge of learning on their own.

# 3. Analysis of Darts Throwing

## 3.1 Introduction

Recently, throwing motions of experts and non-experts have been compared based on biological information such as motion and electromyographic (EMG) signals [70]. For example, Proximal-to-Distal segmental Sequencing (PDS) is found in both joint-angular velocities and EMG signals[71]. PDS indicates such a phenomenon that limb motions are described by successive transitions of a velocity-peaked joint and the beginning of the EMG activation of a muscle from the body trunk to the periphery. PDS can also be observed in gaiting. Finding PDS is attractive because it is strongly related to synergetic motor control and humans can take an optimal motor-control strategy such as minimum-jerk control. On the other hand, the computational property of PDS has not been discussed, though it is widely accepted that humans can take an optimal motor-control strategy such as minimum-jerk control. However, most motor-control experiment were only a reaching movement task between two points, and there was not evaluated in a real exercise skilled motion.In order to understand the human strategy for optimal motor-control, we investigate the motion of dart throwing by quantitatively comparing experts and non-experts based on their scores, motions.

## 3.2 Darts Throwing

We chose dart throwing because it is essentially different from ball throwing in previous studies, as follows. Throwing a dart is one of the discrete skill. The discrete skill is a skill that is organized in such a way that the action is usually brief and has a well-defined begining and end. Discrete skills are tasks such as throwing and kicking a ball, firing arifle, or casting fishing lure. In addition, the dart throwing is the relatively early movement that movement finishes throwing it as 0.5[s]. It is required to learn internal model so that a person realizes such a movement. Because it is thought that the dart throwing is one of the tasks that throwing it needs learning of the internal model in this way, it is thought that we can think about what kind of optimization model it is optimized.

Figure 1 shows an example of throwing motion which captured by high speed

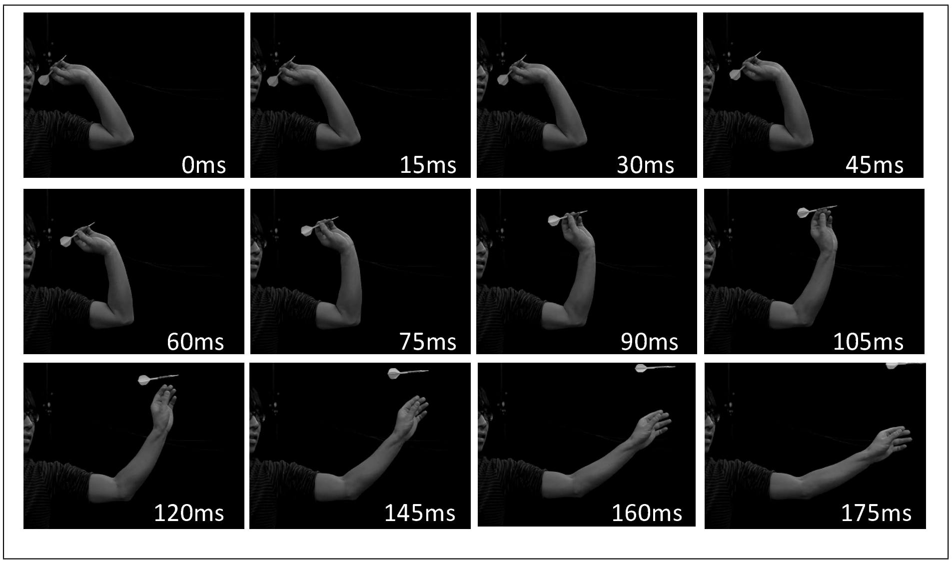



Figure 1. Example darts throwing motion captured by High Speed Camera (1000Hz)

camera (1000Hz). Throwing darts is simple because it usually performed by fixing the body trunk and is primarily driven by an upper-limb(Figure 1). The weight of a dart is much lighter than a ball, and acceleration required in the hand tip for throwing a dart is much smaller than that of a ball. Big acceleration need not be given in the motion of throwing a dart. The possibility of muscle fatigue is much lower in throwing darts. Hence, the influence on the muscle activity caused by fatigue should be much smaller in throwing darts. In contrast to previous studies, we examine the difference between experts and non-experts from the viewpoint of optimal motor control.

## 3.3 Evaluating by Optimization Criteria

### 3.3.1 Optimization Criteria

Each subject's trajectories for each throw were analyzed in terms of the following optimization criteria.

**Sum of squared jerk**

Minimum jerk is an optimization criterion proposed by Flash and Hogan

[72] to explain human motor control. It is known that it precisely explains human reaching movements as long as there is no interaction with external objects. In this study, the arm of each subject did not interact with an external object, except a dart. Because the weight of a dart is much lighter than an arm, throwing trajectories may be well explained by this criterion. The objective function of the minimum-jerk optimization is defined in the task (world) coordinates, and is integration of the squared jerk of a hand for each coordinate during an arm movement.

In this study, we defined an optimization function of the minimum-jerk criterion as

$$C = \frac{1}{2} \sum_{k=t_s}^{t_f} \left( z[k+3] - 3z[k+2] + 3z[k+1] - z[k] \right)^2, \tag{1}$$

where $z$ is coordinate of the hand, $t_s$ is the starting time of a throwing motion and $t_f$ is the ending time.

**Sum of squared joint-torque change**

To overcome the problem of the minimum-jerk criteria which is purely kinematic, minimum torque-change criterion was proposed by Uno, et al. [73] to cover the minimum jerk trajectory model's demerits. The objective function of the minimum torque-change optimization is defined in the joint coordinates, and is integration of squared joint-torque change for each joint during an arm movement. The following is the objective function used in this study.

$$C = \frac{1}{2} \sum_{k=t_s}^{t_f} \left( \tau[k+1] - \tau[k] \right)^2, \tag{2}$$

where $\tau_i$ is joint torque of the $i$th joint, $t_s$ is the starting time of a throwing motion and $t_f$ is the ending time.

Minimize joint Jerk model

This evaluation model is to minimize each joint jerk, was proposed Osu [74].

$$C_{aj} = \frac{1}{2} \int_{t_{start}}^{t_{end}} \left( \frac{d^3\theta_i}{dt^3} \right)^2 dt \tag{3}$$

where, $\theta_i$, $t_{start}$, $t_{end}$ are $i$th joint angle, begining time of take-back, time of releasing the dart, respectively.

The correlation between optimization criteria and score is used for comparison the each criteria.

Table 1. Subjects

| Subject | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| Weight[kg] | 80 | 65 | 61 | 67 | 67 | 80 |
| Height[cm] | 183 | 172 | 172 | 176 | 182 | 171 |

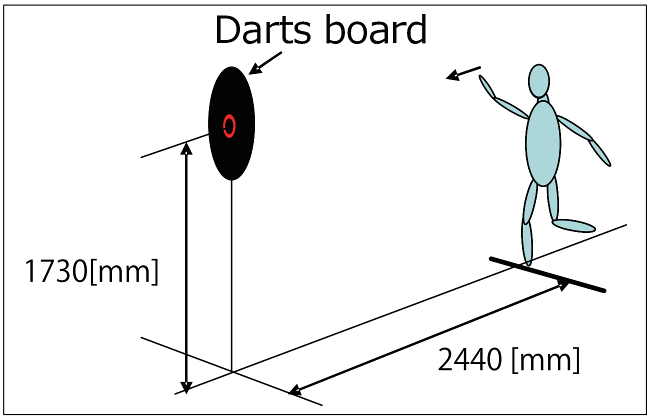



Figure 2. Experimental environment

Table 2. Detail of dart

| Height[mm] | Weight[g] | Flights Shape | Diameter[mm] | Shaft length[mm] |
|---|---|---|---|---|
| 141 | 10 | Standard | Max. 5 | 46 |

## 3.4 Methods

### 3.4.1 Subjects

Six healthy subjects (adult males, age $25 \pm 1$ years ) participated in this experiment. Their body parameters are shown in TABLE 1. We classified them into two groups based on their darts scores.

### 3.4.2 Experimental setup and data preprocessing

The task was soft-tip darts. The goal of this task was to shoot a bull's eye on a dart board. The setting of the dart board and the standing location of

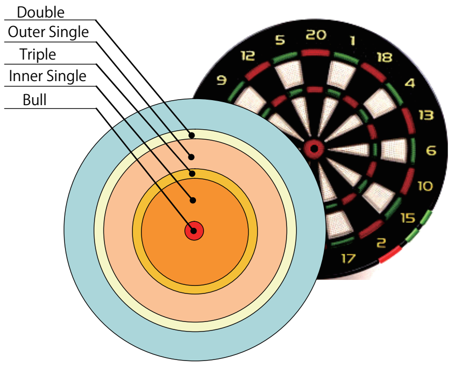



Figure 3. Define the section name of darts-board

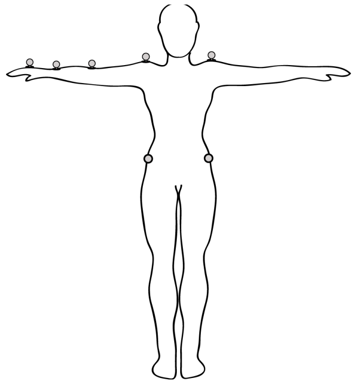


Figure 4. Attached marker positions based on Helen Hays marker set

the subjects followed the official rules of the World Darts Federation (WDF) as shown in Fig. 2. Fig. 3 shows the define of the score. Scores of bull's eye, inner single ring, triple ring, outer single ring and double ring are 5, 4, 3, 2 and 1 point, respectively. Subjects were instructed to shoot for the bull's eye as much as possible with their preferred rhythm. Before the actual task, the subjects were asked to throw darts 30 times. The actual task consisted of 12 trials. In one trial, the subjects initially held four darts with their right hand, and threw them one by one.

We used PC DARTS (Epoch CO., LTD) consisting of a board with a USB connection to a PC, and darts with a soft tip. The scores were automatically calculated by the PC DARTS. We used a MAC3D System (Motion Analysis Corp.) for measuring upper-limb motion.Markers for optical motion measurement were attached to each subject's upper-limb (shoulder, elbow, and hand) according to the Helen Hayes Marker set.

The measured marker positions were low-pass filtered by second order Butterworth filter with a cutoff frequency of 5 Hz. Angular position, angular velocity and angular acceleration of each joint were calculated from the marker positions.

In general, it is said that one throwing motion consists of three phases: the aiming phase, the take-back phase, and the throwing phase. Fig. 6 shows defined these phase on the motion data. We particularly focused on the timing when the aiming phase and the take-back phase was switched, and defined it as the end of

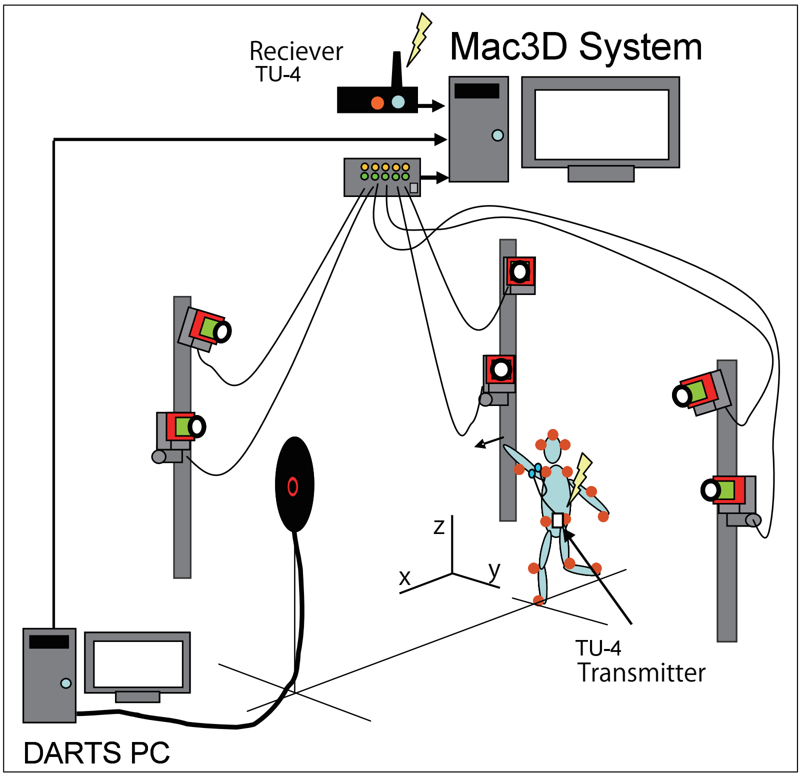



Figure 5. Measurement system for darts throwing

the take-back phase by finding the time when the vertical velocity of the hand tip in the world coordinates became zero. All recorded data were aligned at this switching timing from the take-back phase to the throwing phase.

#### 3.4.3 Estimation of joint torque change

We estimate joint torque that calculates inverse dynamics using the Newton Euler method. Required parameters of mass, center of mass (COM) and inertia were set based on body length and body mass according to [16]. The upper-limb was modeled by three segment mechanical links with five degrees of freedom (DOFs). The shoulder joint was modeled as a 3 DOFs balland-socket, and the elbow and the hand joint were modeled as 1 DOF hinges.

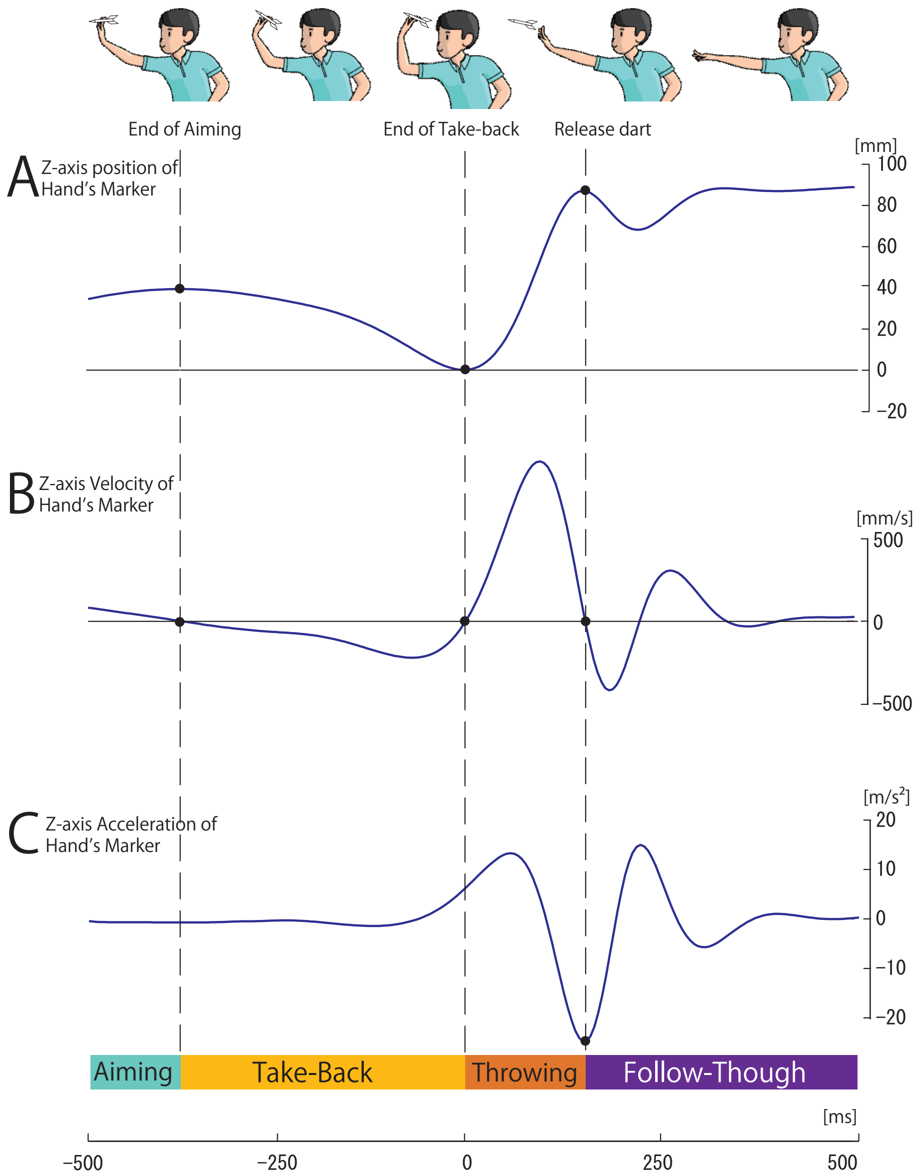



Figure 6. Diagram illustrating how 3 timings of throwing were measured during a throw.

### 3.4.4 Computation method of correlation

Figure 7 shows the computation method of correlation. We use all subjects all data for comparison the property of human.

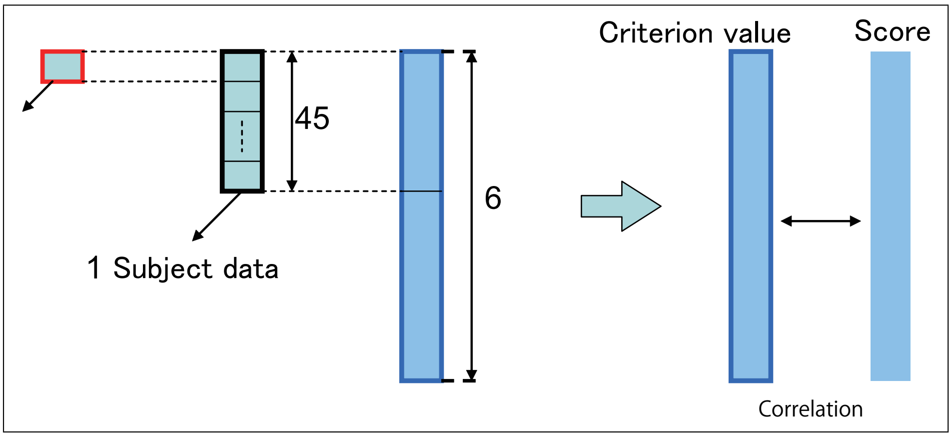



Figure 7. Computation method of correlation between optimization criteria and score

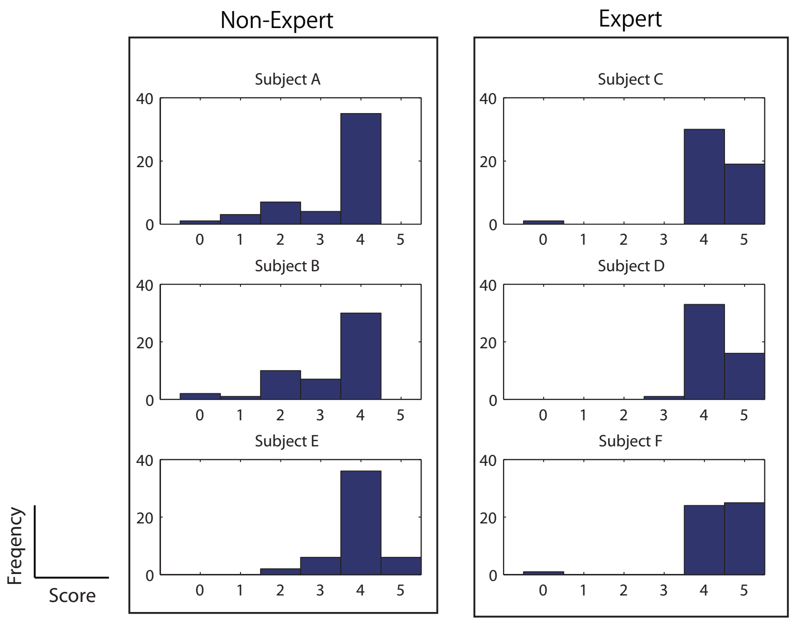



Figure 8. Score distribution of all subjects

### 3.4.5 Results

Figure 8 shows each subject's score. In the results, subjects C, D and F (lower panels) hit the bull's eye better (over 30% of throws) than subjects A, B and E

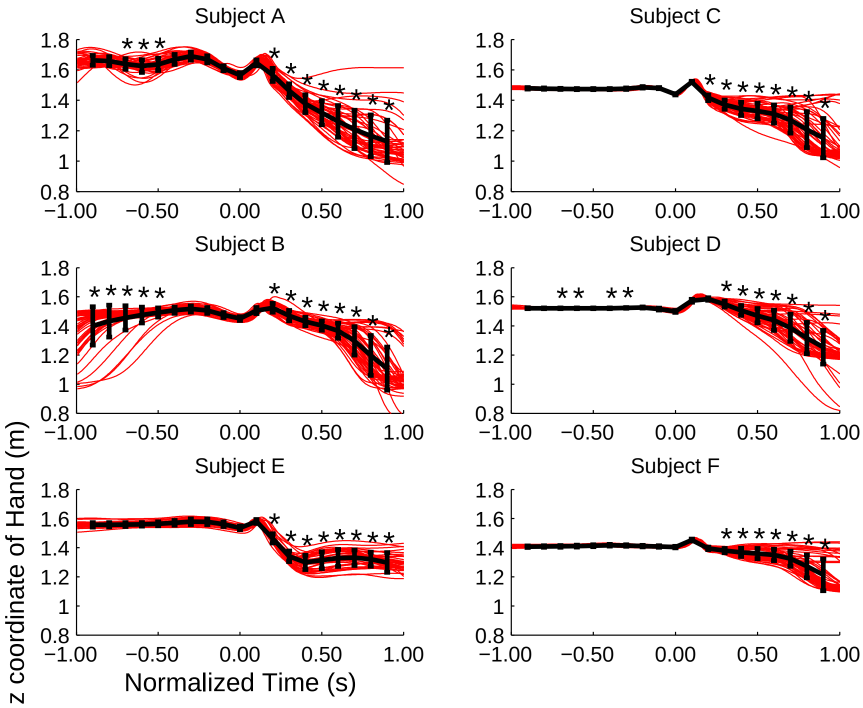



Figure 9. Trajectories of the hand tip of all subjects

(upper panels). Hence, we decided that subjects C, D and F were experts and subjects A, B and E were non-experts for this study.

In Fig. 9, each panel shows whole trajectories of the z-coordinate of each subject's hand during one throw consisting of aiming, take-back, and throwing phases. Hand trajectories are shown by the red lines. Each black error bar shows the variance at a time over trajectories. The left three panels were of non-experts and the right three panels were of experts. In each panel, the horizontal axis is normalized time. '*' indicates the time points at the beginning of the throwing phase, and their variance was shown to be significantly different ($p < 0.05$). This figure clearly shows the significant difference between the experts and the non-experts such that the experts' variance of the hand position in the aiming and take-back phases was much smaller than that of the non-experts.

Figure 10 shows hand jerk trajectories of all subjects. The left three figures were of the non-experts, while the right ones were of the experts. The trajectory

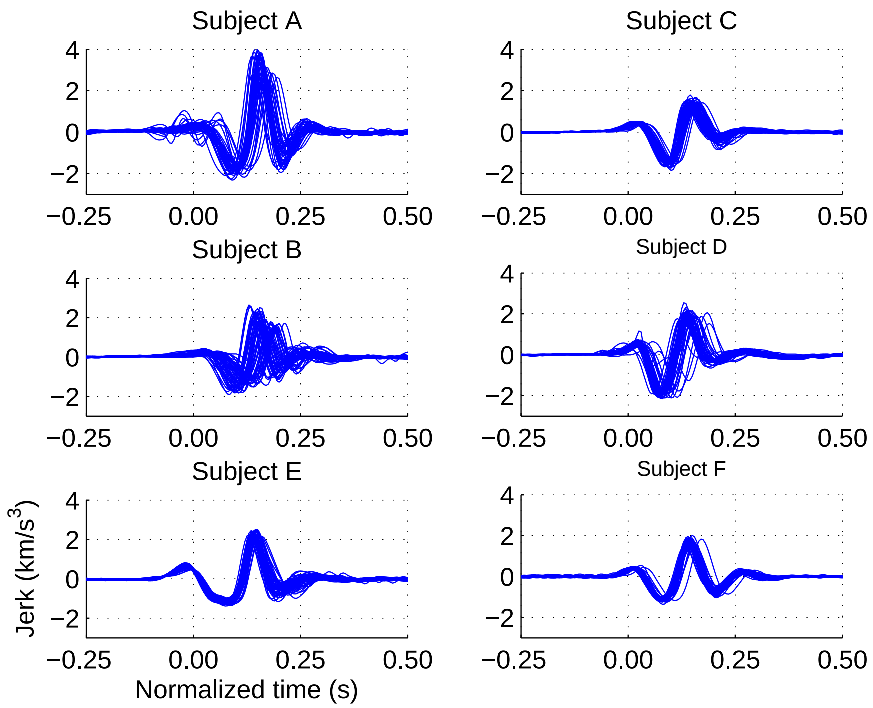
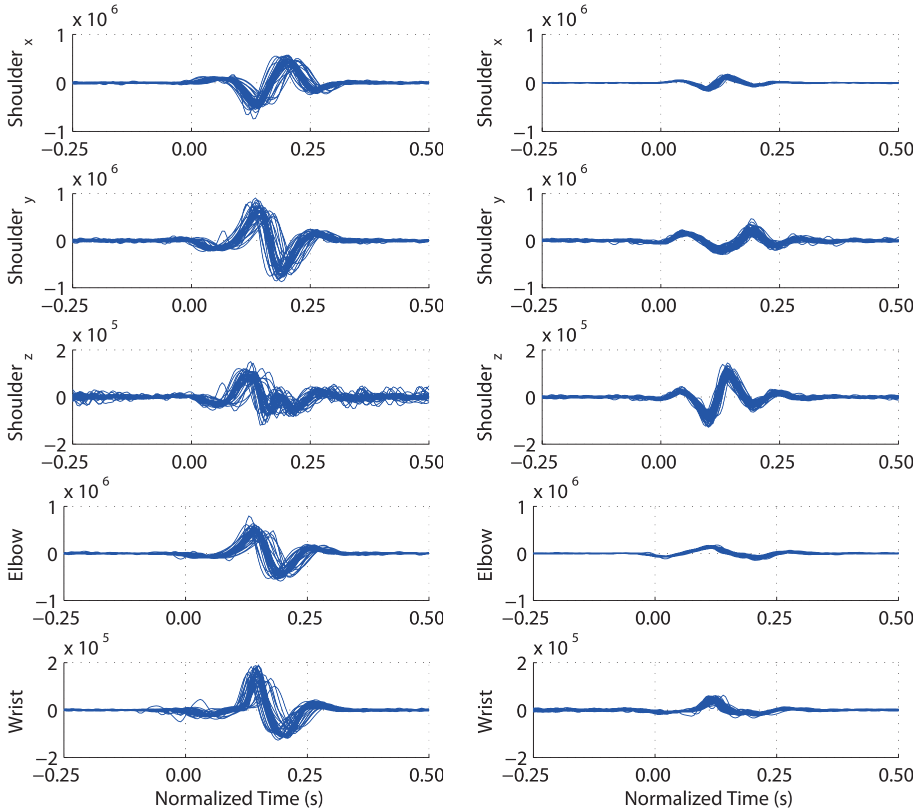



Figure 10. Jerk trajectories of the hand tip of all subjects

amplitude of subject A seems the largest, while both trajectory variances of subjects B and D seem large. The amplitude of vibration and wave pattern were different among subjects. Moreover, significant correlation was not found between the sum of squared jerks and scores of all subjects.

Estimated joint-torque change trajectories of subject F (expert, right panels) and subject A (non-expert, left panels) are shown in Fig. 3.4.5. Five panels in each column correspond to 3 DOFs shoulder-torque trajectories of all throws, 1 DOF elbow-torque trajectories and 1 DOF hand-torque trajectories, respectively. This figure clearly shows that the variance of the non-expert's torque-trajectory was higher than that of the expert. In contrast to the case of the sum of squared jerk, significant correlation was found between the scores of all subjects and their sum of squared torque-change values around the shoulder joint (rotation around x-axis and y-axis), the elbow joint, and the hand joint. The correlation values were -0.26, -0.17 and -0.19 ($p < 0.05$).




(a) Joint torque change trajectory of SubjectA

(b) Joint torque change trajectory of SubjectF

Figure 11. Estimated joint torque change trajectories of an expert (a) and a non-expert (b)

As shown in Fig. 1, subjects stood with their right shoulder forward. With this standing posture, rotation around the x-axis of the shoulder joint corresponds to elevating motion and is caused by shoulder adduction and abduction. Rotation around the y-axis of the shoulder joint also corresponds to elevating motion and is caused by horizontal shoulder flexion and extension. It is reasonable that these two axes of the shoulder joint were elaborately controlled for throwing darts because they mainly contributed to the throwing motion, while rotational arm motion around the z-axis should not. The obtained negative correlation between the sum of squared torque change and the scores of all data suggests that the experts optimally controlled the shoulder elevations, rotation around the elbow, and rotation around the hand joint, in terms of the dynamics, for throwing darts.

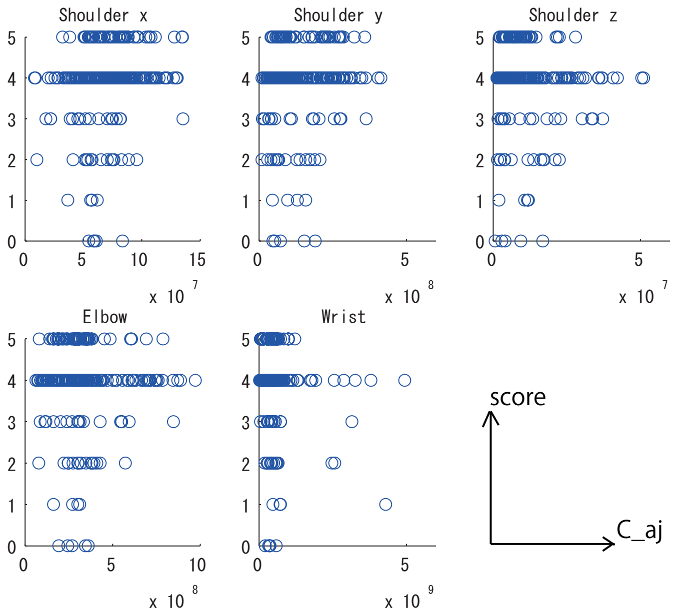



Figure 12. All subject's jerk time series

Fig. 12 shows scatter graph of each joint jerk and score over subjects. Shoulder joint:x,y axis shows positive correlation ($p < 0.05$), hand joint shows also positive correlation ($p < 0.05$).

## 3.5 Summary

The most interesting finding of this study was acquired by analyzing the upper-limb motions of all subjects in terms of trajectory optimization criteria. That is, their sum of squared joint-torque changes was negatively correlated with their scores ($p < 0.05$), whereas their sum of squared jerks was not, suggesting that the experts optimally controlled the shoulder elevations, rotation around the elbow and the hand joint in terms of dynamics. In the joint-jerk, the shoulder joint:x,y axis shows positive correlation ($p < 0.05$), hand joint shows also positive correlation ($p < 0.05$).

## 3.6 Estimating Release Time

### 3.6.1 Motivation

Acquiring skillful movements of experts is a difficult task in many fields. Novices often fail to find out how to improve their skill even by watching the expert's demonstration. Therefore, it is desirable to develop a computer agent that can automatically find quantitative indices of skillful movements that clarify the difference between experts and novices. Although the quantitative indices can be described not only in the motion space, but also in different spaces such as the torque and the muscle space [34], this article focuses on the indices described in the motion space, as it is the most fundamental and important space, and is therefore frequently used for instructors.

On the other hand, teaching skillful movements to novices by experts is also inherently difficult, because unseen activities such as muscle and neural ones generate the movements. Although there are studies aiming at finding the difference in muscle activities [40] and neural activities [75] between experts and novices in many fields, there is little study that applied those findings to actual training assistance. Although the robot essentially assists its trainee in the motion space, the plausibility of the system is shown by the recent results of our application to darts-throwing training. The most important key of our system is that the robot physically, partially, and adaptively restrains the trainee's motion based on the assist-as-needed principle. The assist-as-needed(AAN) principle is well known in the field of physical therapy empirically, and its physiological rationality was indicated at least in the case of rat locomotion [76].

The AAN principle can be intuitively explained: If there is an assist, trainees tend to rely on it, and thus, their training efficiency is generally lower than in the case without the assist. To estimate the skill level of a person for more natural assistance by ANN, we need to collect more data from experts and novices and compare the data on the basis of other features (release time of the dart and so on). Therefore, I developed a measuring environment for estimating the release time. The measurement devices do not disturb the finger motion and sensing. It has the ability to measure the trajectory of the dart simultaneously with the body motion. In this section. I describe the feasibility of our measurement system and detection of the release time of the dart by 2 subjects. We show that our system can use darts throwing comparing task.

### 3.6.2 Release timing

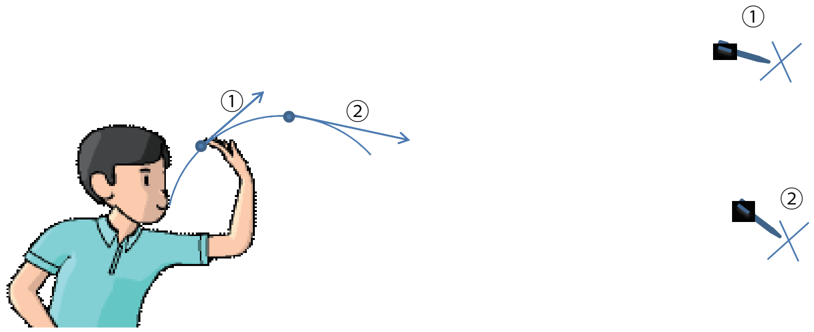



Figure 13. Example of darts throwing

Depending on how the dart is thrown, the score differs. Because, human neural control signals have a noise called signal dependent noise (SDN). We have to control the perturbation to control dart accuracy. Thus, we needed to throw darts at different release timings.

In rigid body space, it is assumed that the dart and hand is one rigid body grasping the dart. We see that a high correlation is shown during the grasping of the dart and highest correlation at release time. We investigated whether the correspondence could be defined by the inner product of the hand and dart velocity, empirically.

### 3.6.3 Experimental setup

The measurement system was prepared based on some conditions: the measurement devices would not disturb the finger motion. The measurement system would have the ability to measure the trajectory of the dart and detect the release time of the dart. Because it is difficult to measure the release time during natural throwing motion by any attached finger devices, we should avoid attaching any devices to the subjects hand. Then the 6[mm] marker is attached to the flights of the dart. The subject's motion with the darts was measured by motion

| | Electrical Switch | Motion Capture of Fingers | Motion Capture of Darts |
|---|---|---|---|
| Subject's Additional Attachment | Two Electrodes and a Transmitter | Ball Markers (2~8) | Nothing |
| Operation of pre process | Attach the electrodes and transmitter | Attach the markers | Nothing |
| Operation of post process | Check the miss detecting | Check the markers of the throwing motion (exchange or miss) | Check the few markers of the dart's motion |

Figure 14. Comparing our method and other studies

capture system this way. Other previous throwing studies did not measure the throwing equipment, quantitatively. It is difficult to perform a number of trials and measure them quantitatively.

**Subjects** Two healthy subjects (age 23, male and female) participated. Eighteen markers, of Helen Hays marker-set, were attached to their body. Their task was to throw 30 darts, aiming for bullseye.

**Experimental environment** In this experimental setup, 9 IR cameras were used and put around the subject and the dart-board to measure the thrown dart. Almost all subjects threw the dart using their right hand. Also, 6 IR cameras were set on the right-side and others set on the left-side to measure the accuracy of the throwing motion and dart.

**Darts** We measured the finger motion for detecting the dart release time. However, it was difficult to detect the marker, which was attached on the finger during throwing. Thus, we chose to measure the dart and attached the small marker on the back of the flight of dart (Fig.16).

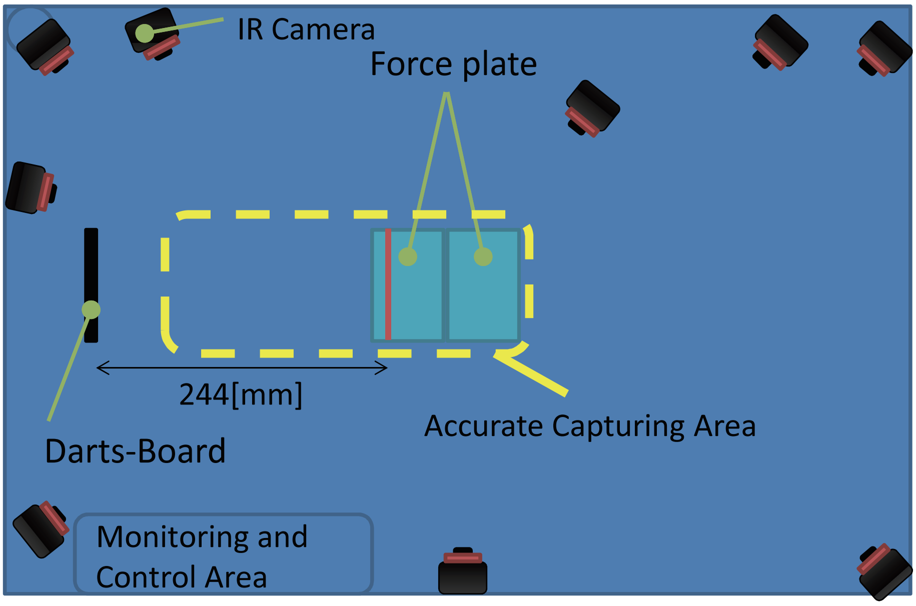


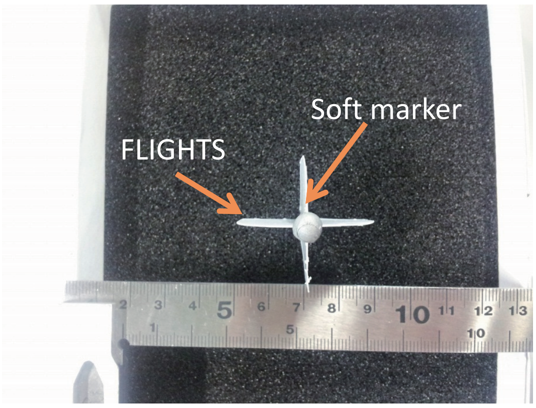

Figure 15. Camera Setting

**Post Data Process** Figure 17 shows the post processing of the motion data. The velocity profile is used to automatically remove outliers using the velocity profile. Further, more outliers were checked and removed manually. Then, linear interpolation was applied. After that, a second-order low-pass filter (cut-off Freqency: 5 Hz) was applied.

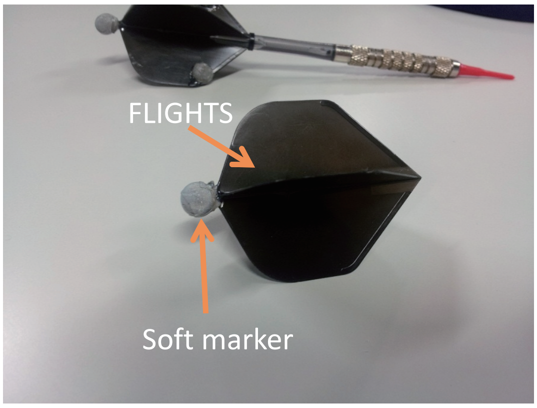
**Similarity of hand velocity vector and dart velocity vector** The similarity refers to that of the inner product of the hand velocity and the dart velocity. In rigid body space, it is assumed that the dart and hand constitute the same rigid body when the dart is grasped. We believed that the high correlation shown during grasping of the dart and the highest correlation would be at release time. the correlation was calculated as follwing:

$$\langle \boldsymbol{v}_{dart}, \boldsymbol{v}_{hand} \rangle \tag{4}$$

where,$\boldsymbol{v}_{dart}$ and $\boldsymbol{v}_{hand}$ are the velocity of the dart's marker and velocity of the hand's marker.




(a) Side view of Flight




(b) Back view of Flight

Figure 16. Flight setting

### 3.6.4 Results

Fig. 18(a) and Fig.18(b) show the trajectory of the hand marker of subjects and the trajectory of dart. Each panels' point (0,0) is the end-time of take-back's hand marker's position.

Figure 19 shows the time change of similarity of each subjects. The boxplot contain each time's similarity of all throws. Figure 19(a) and Figure 19(b) have small variance and high similarity time step. Therefore, we believe that the high

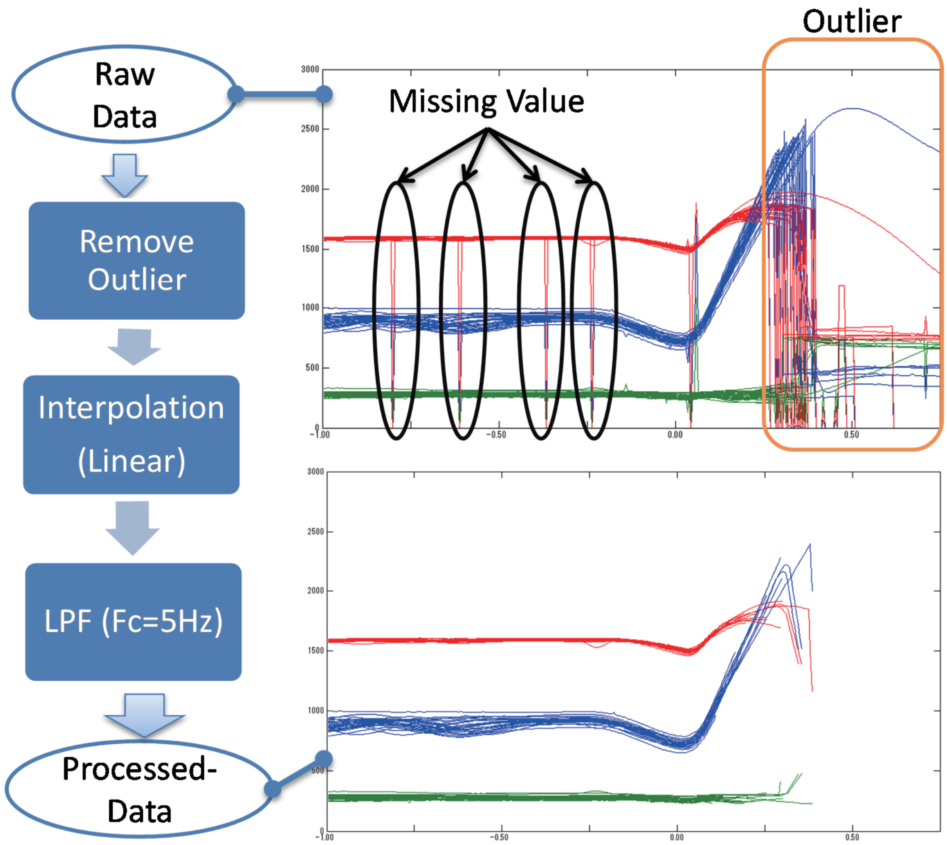
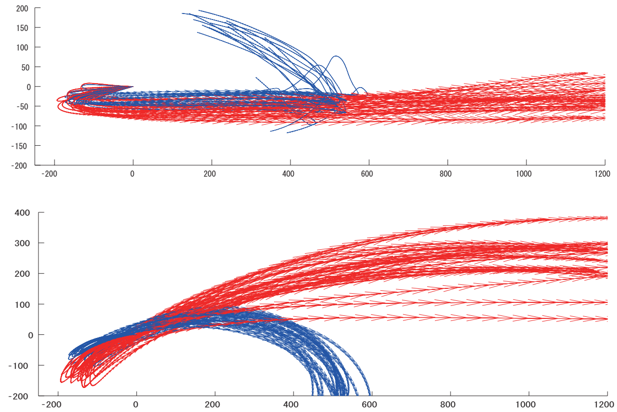



Figure 17. Post Data Process and outlier and missing value examples

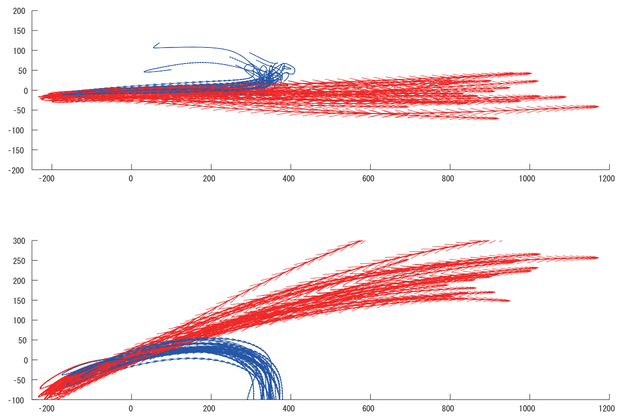
similarity time could be assumed to correspond to the release time.

Fig. 20(a) and Fig.20(b) show the distance between the hand and the dart of each subject, respectively. The end-time of take-back which was defined by [77] is set time at 0[s]. The distance shows monotonic increase after the end-time of take-back. It assumes that the release time can be detected after the end-time of take-back. Therefore, the inner product the hand velocity and the dart velocity after the end-time of take-back were executed. Then, the release time was defined as the peak of the inner product.

Fig. 21(a) and Fig.21(b) show all the release time for all trial, which is the time with respect to the end-time of take-back. The vertical lines are the number of darts thrown by each subjects. The horizontal line is the time that set the



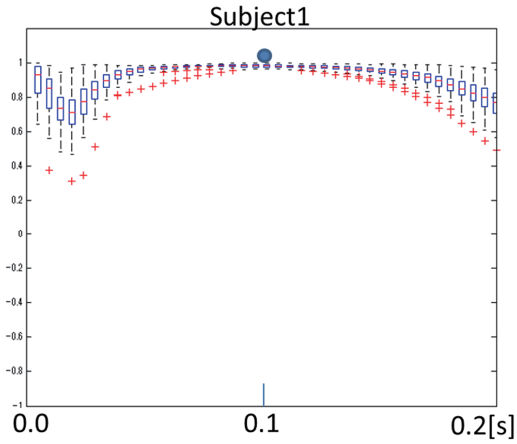

(a) Subject1's hand trajectory and darts trajectory



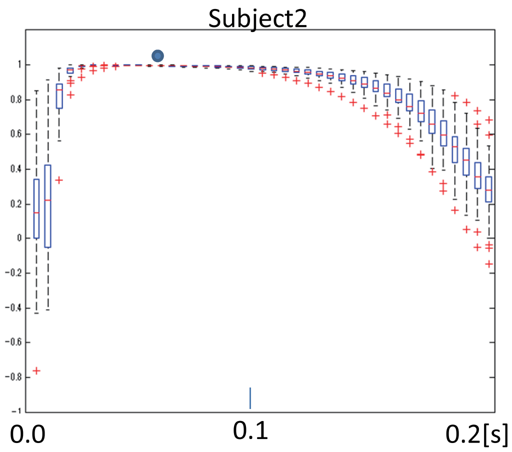

(b) Subject2's hand trajectory and darts trajectory

Figure 18. Subject's hand marker trajectories (Blue Line) and darts marker trajectories (Red Line). Upper panel's vertical line is y-axis(mm) and horizontal line is x-aixs(mm). Bottom panel's vertical line is y-axis(mm) and horizontal line is x-aixs(mm)

end-time of take-back at 0[s]. The variance of subject 1 was larger than that of subject 2. The differences in the release time between subject 1 and 2 are shown



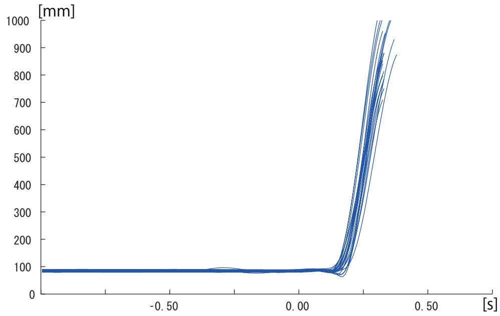

(a) Subject1's time change of similarity



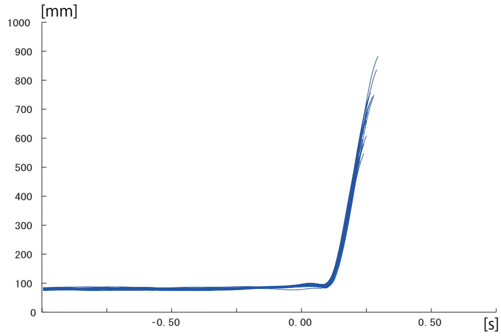

(b) Subject2's time change of similarity

Figure 19. Time change of Similarity of velocity direction of Hand and Darts during throwing phase. Blue circls means estimated release time

by these figures. However, the correlation between the release time and score is not shown.

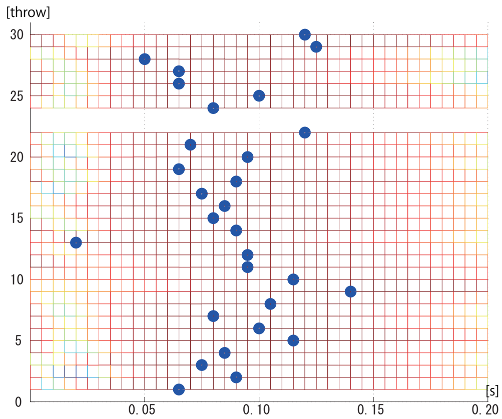



(a) Subject1's Time change of distance of subject1's hand and dart during throwing

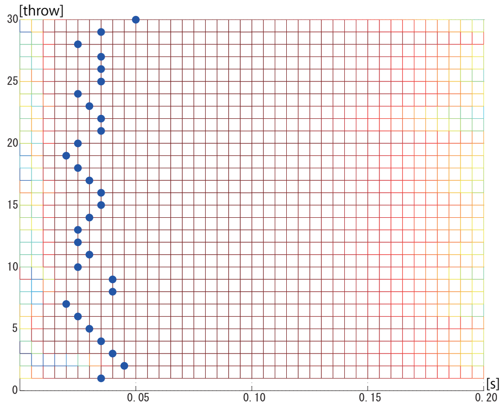



(b) Subject2's Time change of distance of subject2's hand and dart during throwing

Figure 20. Time change of distance of hand and darts during throwing. Vertical line is distance (mm) and horizontal line is time(s)




(a) Subject1's estimated Release Times of 30 throws




(b) Subject2's estimated Release Times of 30 throws

Figure 21. Subjects' release time (circles) of the dart. Vertical line is number of throw (mm) and horizontal line is time(s)

### 3.6.5 Discussion

In this section, the experiment environment to measure a dart was prepared and the pre-experiment was performed. In the results, we suggested that the release time could be estimated by similarity between hand velocity and dart velocity. It is necessary to allow a natural arm throw motion without putting physical burden on the person throwing and to measure the learning degree of progress of the movement. The measurement sensor was small and light and built according to the measurement environment, with IR motion capture system. We used a spherical marker to avoid disturbing the movement of a body and the finger. Using this system, we could measure the movement of throwing an arrow by measuring the movement of the finger of the person of throwing it, but it was necessary to measure the movement of the body limb at the same time, and the number of the markers increases significantly. As for the build, such an individual difference really affects the measurement; every subject has different measurements. It is necessary to regulate the camera layout and focus to prevent occlusion, and a measurement of many arm throw movements becomes difficult in a short time. Even if we could measure the flight by using many small markers, the measured data is lost due to occlusion or the resolution of the camera. Therefore, post-process becomes complex and time-consuming.

Figure 17 shows an example of measured data of a real dart. We understand that a loss is often mixed during throwing movement, even using a few markers. Therefore, we adopted a technique to estimate the time from when I separated an arrow without increasing the number of the markers as much as possible. Specifically, I threw the dart with the IR motion capture system and built a movement and dart orbit simultaneous measurement environment by attaching a small and light marker to the arrow of the dart. We decided to attach a spherical object of diameter 9[mm] to the flight tail of the dart, and a marker to attach the considered mounting location so that influence of the occlusion by finger and flight became small and do not greatly change the flight's aerodynamic characteristics.

Fig. 18(a) and Fig.18(b) show that we can measure the marker trajectory of a dart and the fingers. The two subjects were of different gender, but we could measure the movement of the person and the trajectory of the dart simultaneously. Fig. 21(a) and Fig.21(b) suggest that it is possible to estimate the time a dart separates by this measurement set up. We need to measure the correct release time, and our method will be necessary to verify the accuracy of the data in future.

## 3.7 Summary

In this section, we build a measurement method to estimate the release time of dart throwing. The most interesting part of this study was that the release time was estimated by using only a motion capture system. The measurement environment for the human motion and the dart was prepared by IR motion capture system. The number of the markers and size on the flight of dart were determined experimentally. We suggested that the estimated release time by the dart's position and velocity. We needed to measure the correct release time, and it is necessary to verify the accuracy of the data in future.

# 4. Development of an Adaptive Robotic Trainer for darts throwing

## 4.1 Adaptive Robotic Trainer

The goal of the adaptive robotic trainer for darts throwing was maximizing the subject's score of the darts throwing and minimizing the physical support to avoid learning to keep relying on the physical support.

It is nontrivial for each subject to define the optimal throwing trajectory as well as the optimal policy for the physical support in advance, due to individual differences in the body dynamics and in the neural controller. Therefore, we employed an on-line version of GARB algorithm [78], a policy-gradient based reinforcement learning algorithm, to achieve the goal. An advantage of the policy gradient method is that the policy representation can be chosen so that it is meaningful for the task and can incorporate domain knowledge. This often requires fewer parameters in the learning process than in value-function based methods. Another advantage is that the policy gradient method is a model-free approach. Because of these advantages, it has been applied to robot learning studies including human-robot interaction (HRI) studies [50] [49] [79].

### 4.1.1 Training system overview

This training system (Fig. 22) integrated the following three subsystems through Ethernet: the motion capture subsystem (Mac3D system, Motion Analysis Corp. USA), the dartboard observation subsystem and the manipulator (PA10, Mitsubishi Heavy Industries, Japan) control subsystem. The distance between the dartboard and each subject was 244[cm], and the height of the dartboard from the floor was 173 [cm]. The three-dimensional coordinates of the motion capture subsystem is depicted, and throwing motions are mainly described in the X-Z plane. The experimental setup (Fig. 22) for this learning task was almost the same as the one described in section 3 except that a manipulator was involved. This manipulator (Fig. 23(b)) was employed to physically assist the upper arm of each subject (Fig. 23(b)), but the subject's upper limb was not fixed completely. The manipulator's end-effector was fitted with a soft rubber attachment(Fig. 23(a))

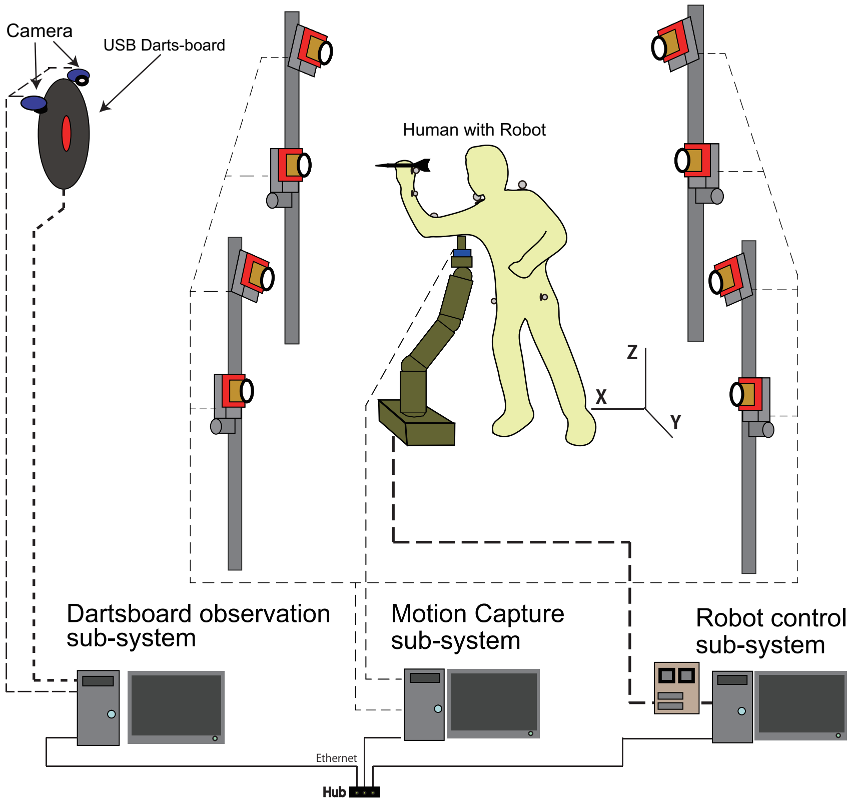
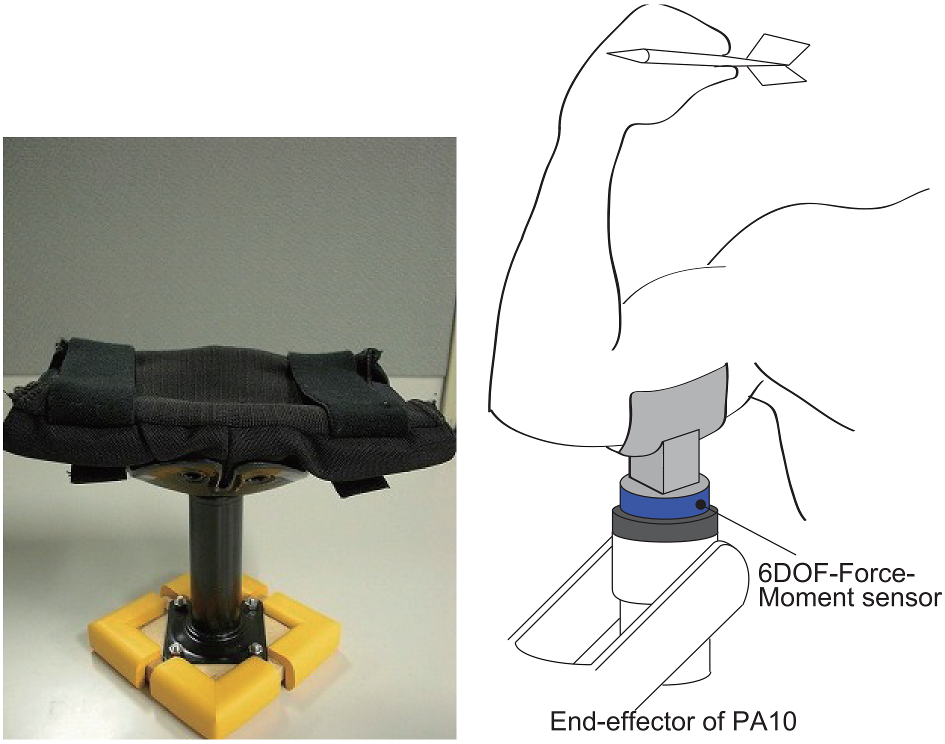



Figure 22. Overview of the training system which integrates the following three subsystems through Ethernet: motion capture subsystem, dartboard observation subsystem and robot control subsystem. The three-dimensional coordinates of the motion capture subsystem is depicted, and throwing motions are mainly described in the X-Z plane.

to partially hold the subject's upper arm safely.

The manipulator's end-effector's initial position was set up based on the following: (1) the subject's upper limb pointed to the dartboard; (2) its posture was level. We chose this setup for the following two reasons. First, we have already shown that elbow displacements $d_e$ and shoulder displacement $d_s$ were larger in novices than in experts. The second reason was for minimum constraint (see Sec. 4). As we have discussed, the state vector for reinforcement learning at the $k$th learning iteration was chosen as $\boldsymbol{s}_k = [d_s, d_e]^T$.

To compliantly assist the subject's upper limb, and to implement the assist-




(a) (b)

Figure 23. Interface to hold a human upper limb. It is made of rubber (a), and attached to the robot as shown in (b)

as-needed function, active impedance control was applied to the manipulator. The active impedance control enabled the manipulator to emulate a spring-mass-damper system with the help of a 6 degree of freedom (DOF) force-moment sensor mounted at the manipulator's end-effector (Fig. 23(b)). The dynamics of the robot's end-effector in the xyz-coordinates are given by

$$M\frac{d^2\boldsymbol{P}(t)}{dt^2} + C\frac{d\boldsymbol{P}(t)}{dt} + K\boldsymbol{P}(t) = \boldsymbol{F}, \tag{5}$$

where $M$, $C$ and $K$ are mass, viscosity and spring (stiffness) of the robot hand. (please see appendix 5) The desired joint angular velocity $\dot{\theta}$ was calculated in real-time through the Jacobian of the robot from $\boldsymbol{P}(t)$ and was sent to the robot

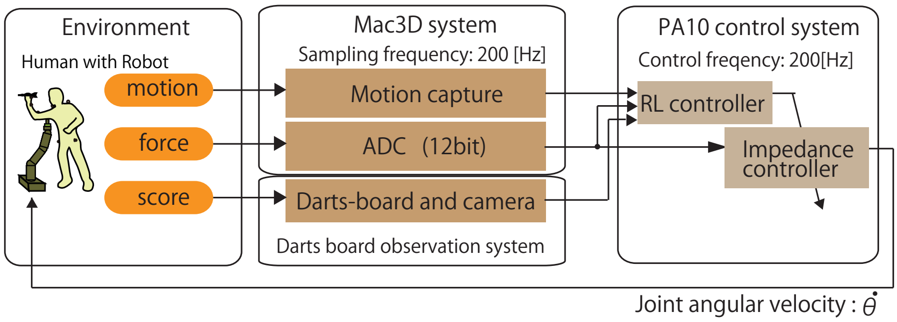



Figure 24. Data flow for the reinforcement learning controller and the impedance controller of the robot

Table 3. Impedance parameters

| Impedance parameter | $K[N/m]$ | $C[N/m^2]$ | $M[kg]$ |
|---|---|---|---|
| Maximum value | 400 | 10 | 2 |
| Minimum value | 0 | 10 | 2 |

controller. The position of the manipulator's end-effector, $\boldsymbol{P}(t)$, was determined by the force $\boldsymbol{F}$ that was measured by the force sensor at a sampling rate of 200 [Hz] (Fig. 24). The control frequency was also the same.

To minimize the strength of the physical assistance according to the increase in subject's skill measured by the score, we employed the online GARB algorithm ([78]), a popular policy-gradient type of reinforcement learning algorithm that has often been used in robot learning studies ([50], [49], and [79]). The policy for the online GARB algorithm in this study was designed to change the stiffness parameter $K$ of the robot's impedance as

$$a_k \sim \mathcal{N}(\cdot; \mu_k, \sigma_k), \tag{6}$$

$$\mu_k = [w_{1,k}, w_{2,k}]\boldsymbol{s}_k, \tag{7}$$

$$\sigma_k = \frac{1}{1+\exp(-w_{3,k})}, \tag{8}$$

$$K_k = K_{def}/(1+\exp(-a_k/K_{def})), \tag{9}$$

where the index $k$ means the $k$th iteration (number of throws) and $K_{def}$ is the default stiffness value. The initial values of the virtual impedance were determined by referring to human hand impedance ([80]). $K_{def}$ was set to a high value, which enabled the upper arm of the subject to be placed reliably on the robot arm. Table 3 shows the maximum and the minimum value of $K$, $K_{def}$, and other initial values for viscosity ($C$) and mass ($M$). $w_i$ ($i = 1, 2, 3$) are the weights to be optimized via the online GARB algorithm.

The reward function was designed to be positively proportional to the score and negatively proportional to the assistive force given by the robot as

$$r_k = \hat{d}_k - \eta \bar{f}_k, \tag{10}$$

where $\eta$, $\hat{d}_k$ and $\bar{f}_k$ are the balancing constant, normalized score and mean assistive force during the throwing, respectively.

$$\hat{d}_k = \frac{d_{max} - d_k}{d_{max}}, \tag{11}$$

$$\bar{f}_k = \frac{f_k}{\Delta t_k}, \tag{12}$$

where $d_k$ is the $k$th distance between the center of the dartboard and the location that a dart hit on the dartboard, and $d_{max}$ is equivalent to the radius of the dartboard. $f_k$ and $\Delta t_k$ are the accumulated assistive force during the $k$th throw, and the duration of the $k$th throw, respectively. Qualitatively speaking, this reward function is designed to be high when the score is high while the assistive force is low. Thus, the goal described by this reward function is to get high scores without the physical assistance of the robot. $\eta$ was empirically set to 0.6. Furthermore, the forgetting factor and the discount factor of learning for the online GARB algorithm were empirically set to 0.92 and 0.98, respectively.

## 4.2 Training experiments

To validate the plausibility and feasibility of the proposed training method, we used three experimental conditions: (1) without robot, (2) with non-adaptive and fixed stiffness (NA-FS) robot, (3) with adaptive robot, and (4) with non-adaptive and decreasing stiffness (NA-DS) robot. In the conditions (2)- (4), we used the same robot. The robot joints were complete fixed and not controlled in the condition (2). Our proposed method was employed in the condition (3). We expected the learning of darts throwing would be facilitated in the condition (3)

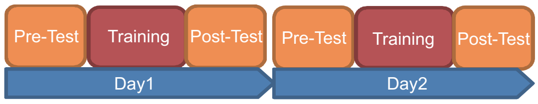
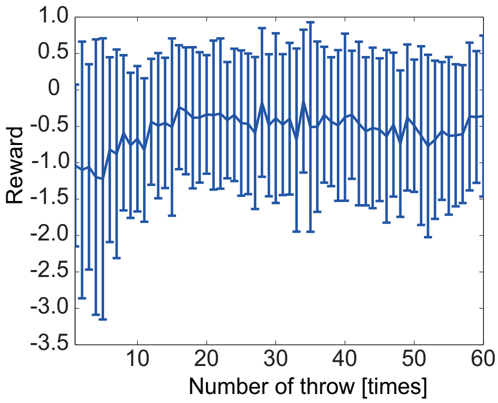



Figure 25. Experimental procedure over two days. In the test blocks, each subject threw darts without the robot in the all conditions so that we can compare the subject's performance after training and before training.

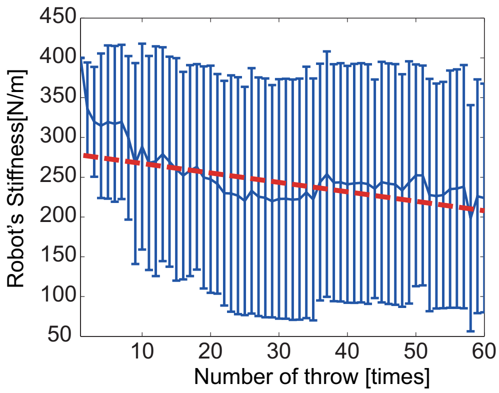
compared not only to (1) but also (2) since the learned skill of the subjects in the condition (2) would heavily rely on the fixed robot. In the condition (3), the initial $K_{def}$value was set to a large value so that it was impossible for subjects to move the robot by their upper limb, which is the same in the condition (2). In the condition (4), the same impedance controller with the condition (3) was used, but its stiffness parameter was constantly decreased over throws. The same constant slope was used in this condition (4). The slope was determined to investigate whether the learned performance in this condition could be comparable to the one in the condition (3). Therefore we determined the slope by linear fitting of the time evolution of the stiffness values obtained in the all experiments in the condition (3) (see the dashed line in Fig. 26(b)). Six different novices among the 24 novices were randomly assigned to each condition. Six different novices among the 18 novices were randomly assigned to each condition. These novices had participated in the experiment that compared experts and novices. Fig.25 shows the experimental procedure over 2 days. Each subject was involved in the first test block (Pre-Test), then in the training block (Training), and finally in the second test block (Post-Test). In the test blocks, each subject threw darts without the robot in all conditions. Thus, we could compare the subject's performance after training with that before training. The numbers of throws in these blocks were 30, 60, and 30, respectively. At the beginning of each training block, the initial height of the robot's end-effector was set according to the subject's preference. The whole procedure was conducted over 2 successive days.

To eliminate some factors affecting motor learning during the experiment ([81]), all subjects read the instructions. They committed to only training dart-throwing during our experiment. They were motivated by the experimenter to improve their performance in the experiments across the 2 days.

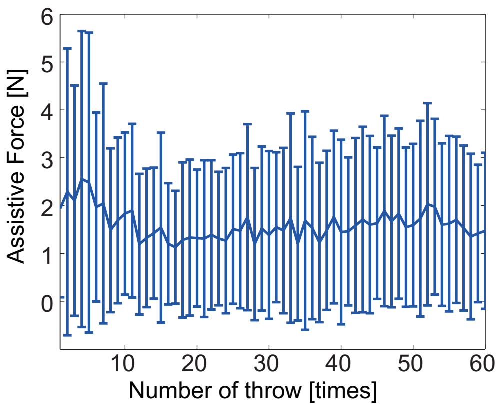
## 4.3 Results




(a) Reward




(b) Stiffness




(c) Assistive force

Figure 26. Time evolution of (a) reward, (b) stiffness, and (c) assistive force in the case "with adaptive robot" condition. Each panel shows the mean and the standard deviation over subjects and days for each throw. Dashed line in (b) shows fitted slope which was used in the case "with NA-DS robot."

Fig. 26 shows the mean and standard deviation of the reinforcement learning variables from six subject in "with adaptive robot" condition over 2 days. Fig. 26(a) shows that the reward increased over the 60 throws. Fig. 26(b) shows that the stiffness of the robot decreased over the 60 throws, according to the acquired rewards in our system. Fig. 26(c) shows that the assistive force $\hat{f}_k$(Eq.(12)), measured at the end-effector of the robot, gradually decreased as the reward increased. This is non-trivial, because the subject had a possibility of laying their elbow on the robot even if the stiffness parameter was set low. Therefore we can conclude that these panels demonstrate the feasibility of our system in that subjects were trained to increase their score and to decrease their dependency on the robotic physical assistance.

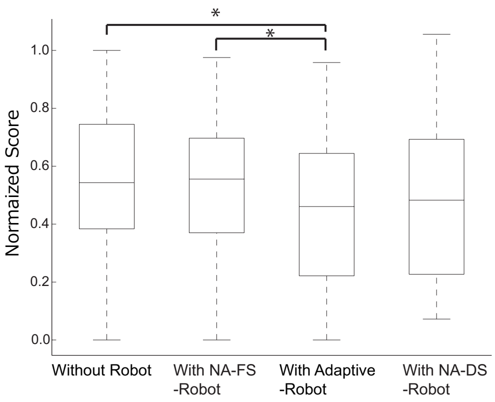
Fig. 27 shows the performance on the 1st day's test. There was a significant difference in normalized score between the "with adaptive robot" condition and the other conditions except the "with NA-DS robot" condition (Tukey-Kramer, $p<0.05$), and also in the shoulder and the elbow displacement among the all conditions (Tukey-Kramer, $p<0.05$). Although the subject group of the "with adaptive robot" condition was more untrained than of the other two conditions, more untrained novices do necessarily improve their skill easily.

Fig. 28 shows the amount of increase in the normalized scores of novices, acquired in the post-test block in the second day and in the pre-test block in the first day, over six subjects in each condition. The mean amount of increase in the "with adaptive robot" condition was significantly higher than in the "without robot" condition as well as the "with NA-DS robot" condition (Steel-Dwass, $p<0.05$). In addition, we found the variance of the increase in normalized score in the "with adaptive robot " condition was significantly smaller than in the "with NA-FS robot "condition (f-test, $p<0.05$), suggesting consistent increase in normalized score was achieved in the "with adaptive robot" condition.

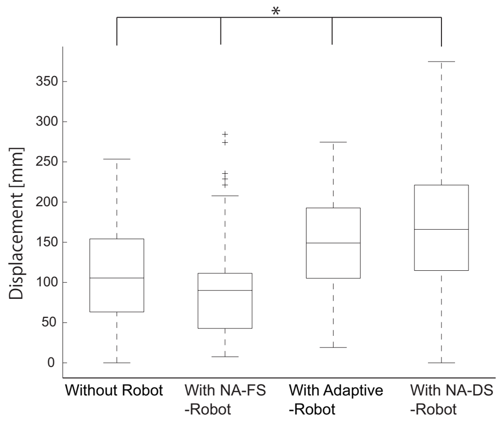
Fig. 29 shows distributions of the daily increase in normalized score in each condition. The mean increase in normalized score of the second day in "with adaptive robot" condition was significantly larger than in "without robot" condition and "with NA-DS robot" condition.

Fig. 30 shows the normalized score distributions in each condition of the final post-test phase. The mean normalized scores in the "with NA-FS robot" condition and "with adaptive robot" condition were significantly higher than in the "with NA-DS robot" condition (Steel-Dwass, $p< 0.05$).

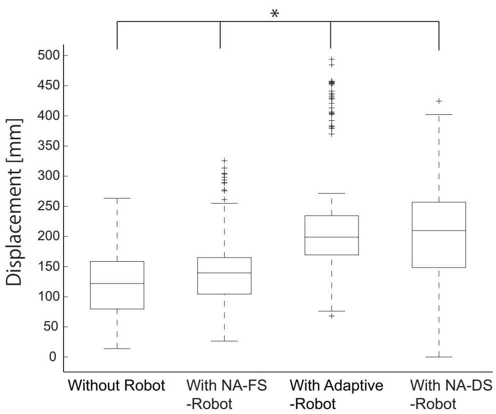
Fig. 31(a) and Fig. 31(b) show the amount of change in the mean shoulder




(a) Score




(b) Shoulder Displacement




(c) Elbow Displacement

Figure 27. Distributions of (a) score, (b) shoulder displacement, and (c) elbow displacement in the pre-test on the 1st day for each condition. These data for all subject in the group were used for plotting each group's distribution.

displacement and elbow displacement between the post-test on the second day and the pre-test on the first day in each condition. As shown in this figure, the mean changes in both shoulder and elbow in the three experimental conditions with the robot used were significantly smaller than "without robot" condition (Steel-Dwass, $p<0.05$).

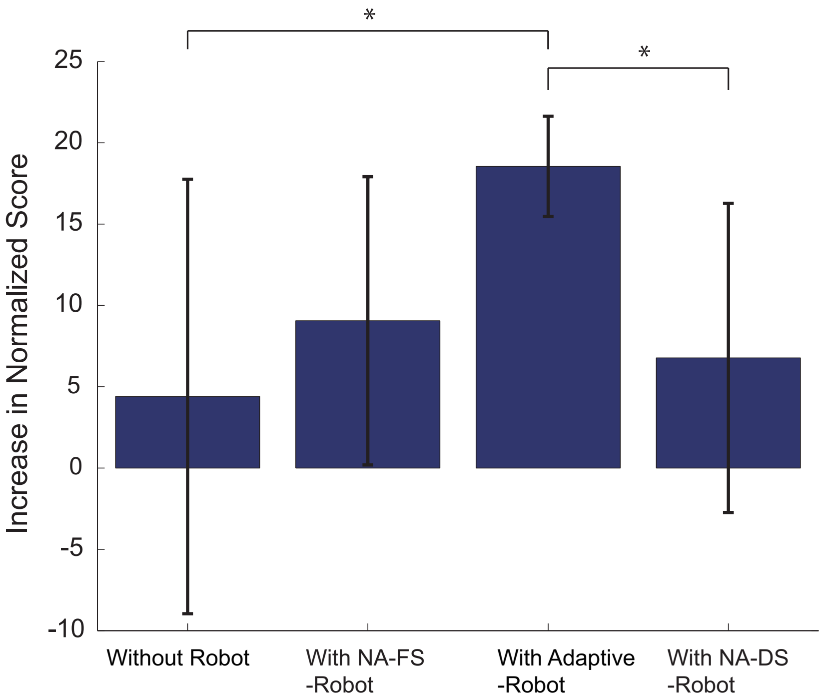



Figure 28. Increase in normalized score in each condition. Each bar and error-bar stand for the mean and the standard deviation over the subjects and the days in each condition.

Fig. 32(a) and Fig. 32(b) show the amount of change of the elbow displacement and the shoulder displacement on each day. The displacements tended to increase in the "without robot" condition and to decrease in the other conditions.

## 4.4 Discussion

### Learning results

- Subjects performance

  The results presented in the last section together suggest the plausibility of our adaptive training system. The normalized score was significantly higher than in the "without robot" condition as well as the "with NA-DS

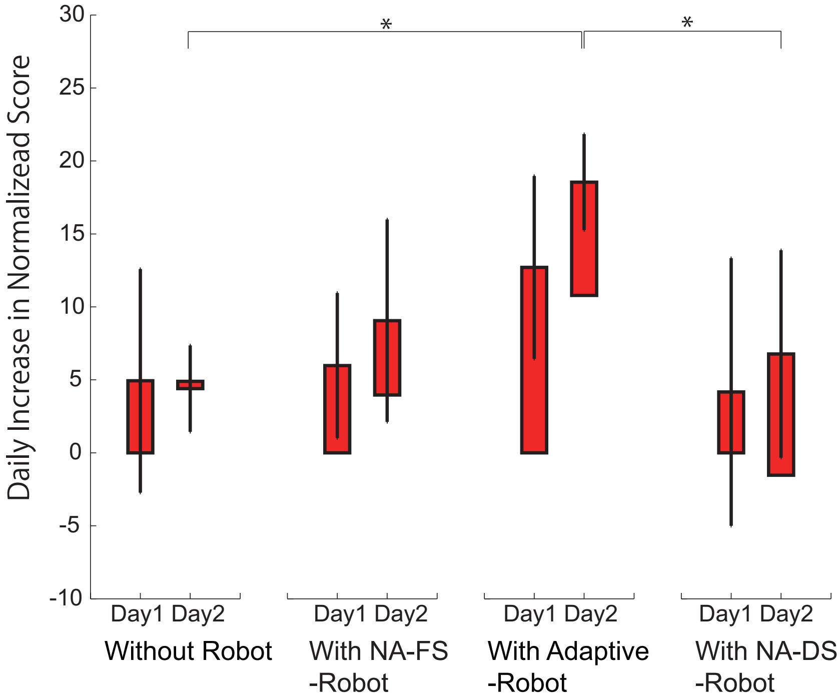



Figure 29. Distributions of the daily increase in normalized score in each condition. Each bar and error-bar indicate the mean and standard deviation over the subjects and the day in each condition.

robot" condition(Fig. 28), which is congruent with our expectation. Although Fig. 27(a) shows the subjects of the "with adaptive robot" condition were more untrained than of the other two conditions, this should not the only reason why we obtained the result shown in Fig. 28. The final normalized score in the "with adaptive robot" condition, indeed, became equal or higher than in the other conditions (Fig. 30) despite it was the lowest before learning. Increase in normalized score in "with NA-DS robot" condition is significantly lower than "with adaptive robot" condition (Fig. 28-Fig. 30). These facts suggest the validity of our implementation of user-adaptation (assist-as-needed) to accelerate learning.

- Three Learning parameters How learning progressed (learning profile) in "with adaptive robot" condition has been shown by Fig. 26. It seems that learning progress was rapidly at the beginning and almost got saturated

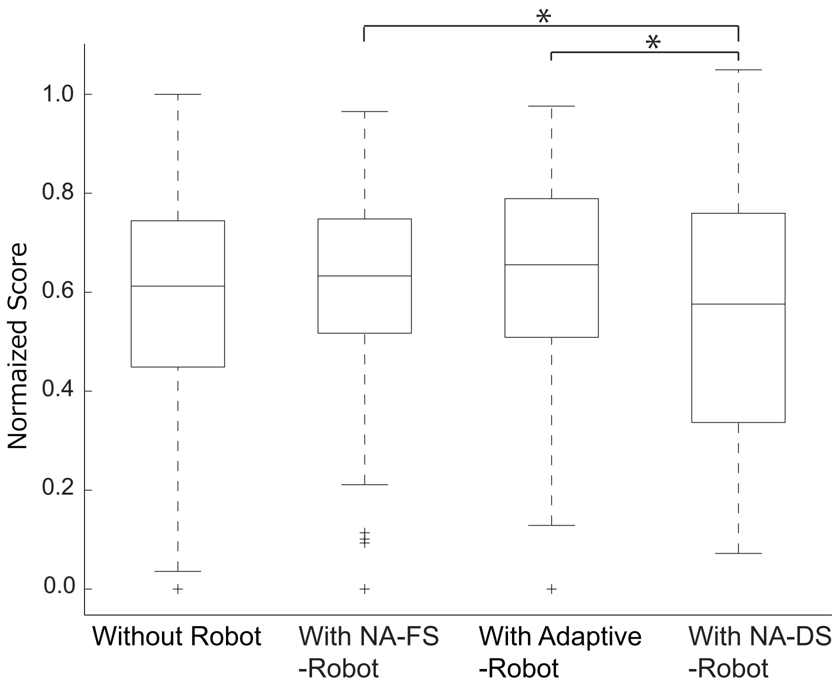
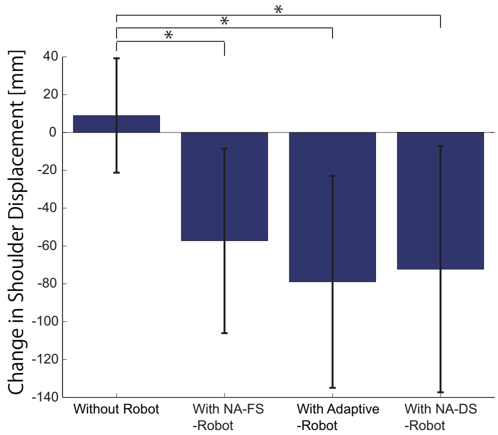


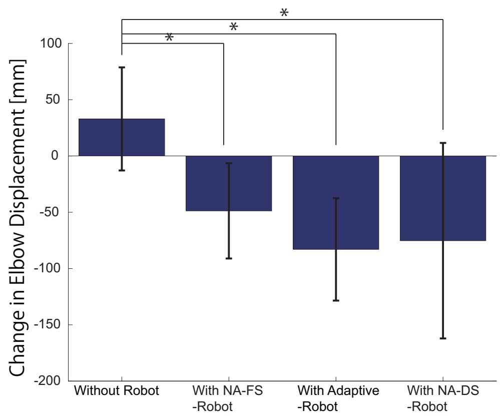

Figure 30. Distributions of normalized scores in each condition of the final post-test. Normalized scores for all subject in the group were used for plotting each group's distribution.

in 10 to 20 throws, which is congruent with the profile of the three weight parameters shown in Fig. 33. The three weights started increasing after 10 to 20 throws, which is because the actual stiffness was saturated although the learning controller kept trying to set a higher stiffness to the impedance controller. The large variances observed in these panels reflect the subjects' individual variability in their training.

- Displacements Fig. 31 and Fig. 32 suggest that the elbow and the shoulder positions got stable when the subjects were involved in the experiments using the robot. Note that stabilizing their positions was not the explicit goal of our training system, but we also had expected this outcome since it is what experts in darts throwing do.




(a)

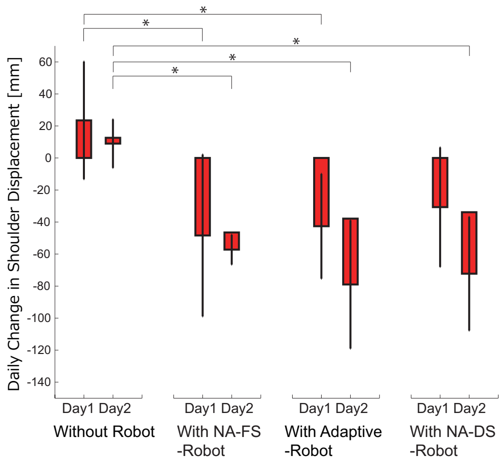



(b)

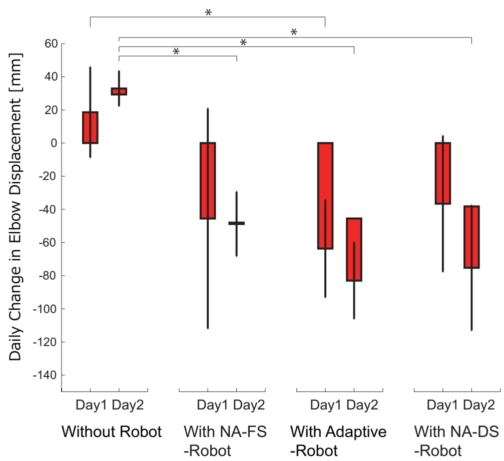
Figure 31. Change in shoulder (a) and elbow (b) displacement over two days in the test block for each condition. Each bar and error-bar indicate the mean and the standard deviation over the subjects in each condition.

**Factors for learning** There are three factors for learning used in this study (see Sec.4.1.1). First, Eq. (10) is the reward function which directly describes our implementation of the assist-as-needed principle. $\eta$ controls the penalty given by the mean assistive force, relative to the normalized score. We empirically set this value to 0.6, through pilot experiments, to see increase in the average reward and decrease in the mean assistive force simultaneously in tens of trials. Second, we also needed to set the forgetting factor and the discount factor for using the online GARB algorithm. The forgetting factor was set to 0.92, a greater value than the average of the human forgetting factor value of 0.76 estimated by [60], in which the "assist-as-needed" controller attempted to reduce its assistance, but at a rate slower than that of the average unimpaired learning human. The discount factor was set to 0.98 which is near 1 and shown to be effective in [78] and [79]

**Differences with other model-free and assist-as-needed training methods** In this study, we have proposed a model-free and assist-as-needed training method for learning assistance, and have shown its feasibility and plausibility




(a)




(b)

Figure 32. Daily change in shoulder (a) and elbow (b) displacement in the test block in each condition. Each bar and error-bar indicate the mean and the standard deviation over the subjects in each condition.

through the learning experiments on darts throwing with healthy subjects. There are also model-free and assist-as-needed training methods designed for rehabilitation, e.g., [82], [46], and [47], but they mostly required the predetermined desired trajectory. Note that it is nontrivial to define the optimal trajectory for each individual in advance, owing to their differences in body kinematics/dynamics and in the neural controller. The aim of our proposed framework is task-goal oriented, instead of the desired-trajectory oriented. Recently [83] proposed a Bayesian approach to robot-assisted motor skill learning and rehabilitation, but their approach does not maximize the expected task-achievement. The key feature of their approach is to develop the direct mapping between the performance and the assistance. In contrast to our method, it is therefore difficult to utilize human domain knowledge as state variables and a policy function to accelerate learning.

**Toward neuro-rehabilitation** As described in section 4 and as discussed above, the framework proposed in this paper is general and can also be applied to

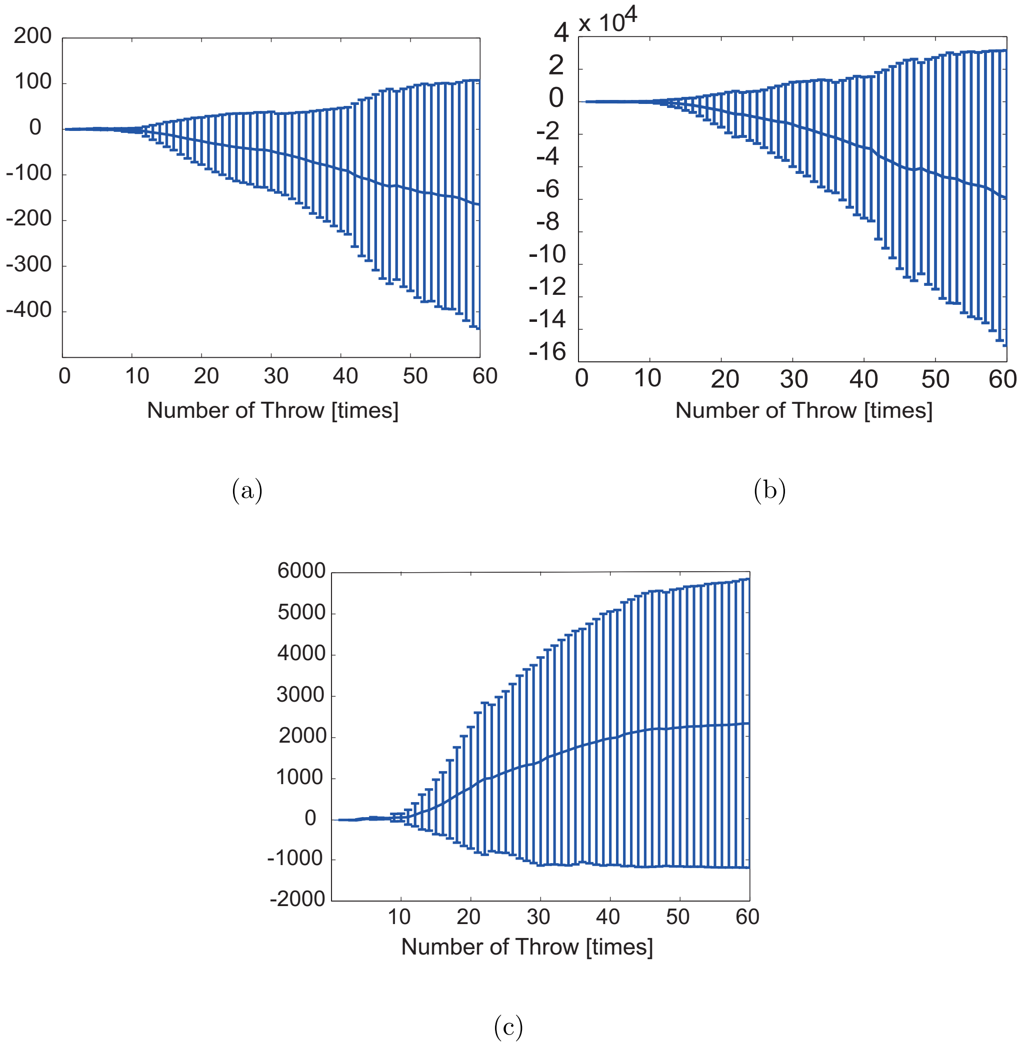



Figure 33. Time evolution of learning parameters (a) $w_1$, (b) $w_2$ and (c) $w_3$ in the case "with adaptive robot" condition. Each panel shows the mean and the standard deviation over subjects and days for each throw.

rehabilitation. Since the central engine in our framework is reinforcement learning, the critical point for applications is whether we can specify (i) state variables, (ii) policy function, and (ii) reward function. Note that they should be specified in a low dimensional space for efficient assistance. One natural application would be to gait rehabilitation for patients with spinal cord injury (SCI), like [46]. They proposed to apply the iterative learning control to the robotic gait rehabilitation, but no results with patients have been shown yet, which might be caused by the fact that they required the desired trajectory to be followed by the impaired limb. Instead of using the desired trajectory, we would use such goals as walking speed and walking distance. Biological signals such as electromyogram (EMG) and cortical activities measured by near infrared spectroscopy (NIRS) could be used to describe the state variables and the reward function.

# 5. Conclusion

In this thesis, I have proposed an adaptive robotic training framework. In addition, I have developed a robotic training system for novices of dart throwing, based on proposed framework. An experiment was setup to investigate the feasibility of the proposed training system.

Section 2 proposed a new framework to realizing non-trajectory based adaptive robotic assistive robotic system. Section 3 described the comparison between an expert and a novice in dart throwing from the aspects of some optimization criteria for motor skill and the features of the measurement environment to estimate the release timing. Section 4 described the development of the proposed adaptive learning assistive robotic system for novice in dart throwing.

How a robot can physically assist a novice was determined based on motion comparisons between a novice and an expert. The novice tends to show large displacement compared with an expert. Since motion comparisons revealed that a novice had larger displacement on their shoulder and elbow during throwing compared to an expert, a robot was used to give assistive force to the upper limb of the novice. Our assist approach was based on four properties: goal-oriented, assist-as-needed, model-free, minimum constraint. Specifically, we allowed for variability in the subject's throwing motion and prevented the subject from learning to rely on the physical assist. The subject's upper-limb was not fixed completely, but was adaptively assisted by a robot control employing a policy gradient type learning algorithm with the aim of maximizing the score of the darts throwing and minimizing the robotic physical assist. Four conditions, i.e., "without robot", "with NA-FS robot", "with NA-DS robot", and "with adaptive robot" conditions, were used to validate the plausibility of the developed system. After screening experiment, 24 healthy novices participated under three types of conditions. In the experimental results, only the score for "with adaptive robot" condition showed statistically significant improvement over 2 days' of training. These results suggest that the subjects in the "with NA-FS robot" and "with NA-DS robot" condition did not increase their score compared with those in the "with adaptive robot" condition. Furthermore, the displacement decreased in "with adaptive robot" condition despite the fact that subjects had minimal assistance (i.e., without strong correction). These results demonstrate that the proposed system is feasible. Under the theme of learning to throw darts, without the track-based assist with a restraint, in this study, we show the leaner could learn dart throwing under minimal restraint and adaptive assist in the proposed framework. Further, the assist robots can also show an example of learning aid to encourage motor learning of a person. The proposed framework performs a task aided design and adaptation algorithms goal-oriented. I believe that it is capable of application in the rehabilitation of stroke patients, such as in existing applications like track-based rehabilitation. The proposed framework requires a quantitative index for evaluating the movement task. However, quantitative indicators exist in rehabilitation and sports fields. Thus, for an adaptive rehabilitation system based on the proposed scheme, we can construct a training system to apply adaptive assist to a person using a quantitative index correlated strongly with training and evaluate the learning state of a learner. It is not necessary to perform learning of the track base by constraining the person. The assistive robotic system has the flexibility of operation for a person's learning and improves safety. The existing paradigm of adaptive rehabilitation robot has been used mainly in the plane of movement at a precise force field production environment. In this case, the person is bound firmly to the assistive robotic system. In our method, there is no need to restrain the person as long as there is no fatal reason to do so, such as for fall prevention. It also does not require a dedicated training to use. We suggest that a humanoid robot can also be used for motor learning assist. Our assist framework was based on repetitive motor learning method that is employed in the training experiments of many motor learning studies. To use the framework in rehabilitation and sport, it is recommended that the training system is built considering psychological aspects such as attention and motivation. Some cases have been implemented in robotic-assisted system considering psychological aspects. However, it was not to perform smooth communication such as by a therapist. The case of increasing the motivation of the person by introducing games have also been shown in recent years. In rehabilitation, I believe that the rehabilitation system allows a comprehensive assist more flexibility by building a system in conjunction with doctors and therapists. In sports training, I believe a more niche learning assist can be expected by designing the system in cooperation with the instructors nor the athlete. Next, my work will demonstrate a proposed assistive control framework application for rehabilitation.

# Acknowledgements

Foremost, I would like to express my sincere gratitude to my supervisor, Professor Kazushi Ikeda for providing me this precious study opportunity as a Ph.D student in his laboratory. I could not have imagined having a better advisor and mentor for my Ph.D study.

I especially would like to express my deepest appreciation to my co-supervisor, Visiting Professor. Tomohiro Shibata for his elaborated guidance, considerable encouragement and invaluable discussion that make my research of great achievement and my study life unforgettable.

I appreciate the feedback offered by my co-supervisor, Professor Tsukasa Ogasawara. I am also very grateful to Dr. Tomoya Tamei for their valuable cooperation in my experiments.

I thank my colleagues in NAIST, Mathmatical inforomatics Lab: Atsushi Miyamoto, Tomoyasu Horikawa, Yusuke Suzuki, Yoshiro Fukui, Akihiro Nakamura and many other students, for the stimulating discussions, for the sleepless nights we were working together before deadlines, and for all the fun we have had in the last a few years. Also, I thank my friends in Kyoto University, Satoshi Nishida.

# References


[1] H.I. Krebs, N. Hogan, M.L. Aisen, and B.T. Volpe. Robot-aided neurorehabilitation. *Rehabilitation Engineering, IEEE Transactions on*, Vol. 6, No. 1, pp. 75–87, 1998.

[2] P.S. Lum, C.G. Burgar, and P.C. Shor. Evidence for improved muscle activation patterns after retraining of reaching movements with the mime robotic system in subjects with post-stroke hemiparesis. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, Vol. 12, No. 2, pp. 186–194, 2004.

[3] D.J. Reinkensmeyer, L.E. Kahn, M. Averbuch, A. McKenna-Cole, B.D. Schmit, and W. Z. Rymer. Understanding and treating arm movement impairment after chronic brain injury: progress with the arm guide. *Journal of rehabilitation research and development*, Vol. 37, No. 6, pp. 653–662, 2000.

[4] D.J. Reinkensmeyer, C.D. Takahashi, W. K. Timoszyk, A.N. Reinkensmeyer, and L.E. Kahn. Design of robot assistance for arm movement therapy following stroke. *Advanced robotics*, Vol. 14, No. 7, pp. 625–637, 2001.

[5] S. Jezernik, G. Colombo, T. Keller, H. Frueh, and M. Morari. Robotic orthosis lokomat: A rehabilitation and research tool. *Neuromodulation: Technology at the Neural Interface*, Vol. 6, No. 2, pp. 108–115, 2003.

[6] H.I. Krebs, J.J. Palazzolo, L. Dipietro, M. Ferraro, J. Krol, K. Rannekleiv, BT Volpe, and N. Hogan. Rehabilitation robotics: Performance-based progressive robot-assisted therapy. *Autonomous Robots*, Vol. 15, No. 1, pp. 7–20, 2003.

[7] Mitsuo Kawato. Computational schemes and neural network models for formation and control of multijoint arm trajectory. *Neural networks for control*, pp. 197–228, 1990.

[8] M.M. Merzenich, R.J. Nelson, M.P. Stryker, M.S. Cynader, A. Schoppmann, and J.M. Zook. Somatosensory cortical map changes following digit ampu-

tation in adult monkeys. *Journal of comparative neurology*, Vol. 224, No. 4, pp. 591–605, 1984.

[9] G.E. Schatz, P.P. Lowry II, and A. Ramisamihantanirina. Changed perceptions in braille readers. *Comptes Rendus de l Académie des Sciences, sér. 3*, Vol. 316, pp. 725–729, 1993.

[10] J. Emken, R. Benitez, and D.J. Reinkensmeyer. Human-robot cooperative movement training: Learning a novel sensory motor transformation during walking with robotic assistance-as-needed. *Journal of NeuroEngineering and Rehabilitation*, Vol. 4, No. 1, pp. 8–15, 2007.

[11] K. Sasaki, H. Gemba, et al. Plasticity of cortical function related to voluntary movement motor learning and compensation following brain dysfunction. *Acta neurochirurgica. Supplementum*, Vol. 41, p. 18, 1987.

[12] Tatsuya N., Tetsuya O., Jun T., Toru T., and Hiroshi O. 3d-3 tool-body assimilation model using recurrent neural network for robot to move objects on a table. *Proceedings of the 71th National Convention of IPSJ*, Vol. 3, p. 3.

[13] Cota N. and Yasuo K. Computational model of tool-body assimilation based on object identification. *Journal of the robotics society of Japan*, Vol. 25, No. 6, pp. 897–905, 2007.

[14] Reza Shadmehr and Ferdinando A Mussa-Ivaldi. Adaptive representation of dynamics during learning of a motor task. *The Journal of Neuroscience*, Vol. 14, No. 5, pp. 3208–3224, 1994.

[15] R.A. Schmidt and C.A. Wrisberg. *Motor learning and performance: a situation-based learning approach.* Human Kinetics Publishers, 2008.

[16] V. Zatsiorsky and V. Seluyanov. The mass and inertia characteristics of the main segments of the human body. *Biomechanics VIII-B*, Vol. 56, No. 2, pp. 1152–1159, 1983.

[17] Michiyoshi A., Hai peng T., and Takashi Y. Estimation of inertia properties of the body segments in japanese athletes. *Journal of the Society of Biomechanisms*, No. 11, pp. 23–33, may 1992.

[18] PF Groshaw. Hardiman i arm test, hardiman i prototype. *General Electric Report S-70-1019, General Electric Co., Schenectady, NY*, 1969.

[19] BJ Makinson. Research and development prototype for machine augmentation of human strength and endurance. hardiman i project. Technical report, DTIC Document, 1971.

[20] Partner robot family. `http://www.toyota-global.com/innovation/partner_robot/family_2.html`.

[21] Stride management assist device to be featured in program to assess elder independence solutions. `http://world.honda.com/news/2012/c120729Stride-Management-Assist-Device/index.html`.

[22] S. Lee and Y. Sankai. Power assist control for walking aid with hal-3 based on emg and impedance adjustment around knee joint. In *Intelligent Robots and Systems, 2002. IEEE/RSJ International Conference on*, Vol. 2, pp. 1499–1504. IEEE, 2002.

[23] Conor James Walsh, Ken Endo, and Hugh Herr. A quasi-passive leg exoskeleton for load-carrying augmentation. *International Journal of Humanoid Robotics*, Vol. 4, No. 03, pp. 487–506, 2007.

[24] C.J. Walsh. Biomimetic design of an under-actuated leg exoskeleton for load-carrying augmentation. Technical report, DTIC Document, 2006.

[25] T. Hayashi, H. Kawamoto, and Sankai Y. Control method of robot suit hal working as opeartor ’ s muscle using biological and dynamical information. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2005)*, pp. 2–6. IEEE, 2005.

[26] Shigeo T. and Junichiro Y. Development of the powered assisted suit system(¡special issue¿development of user-oriented human-support robots). *Journal of Precision Engineering*, Vol. 73, No. 3, pp. 305–308, mar 2007.

[27] E. Yagi, D. Harada, and M. Kobayashi. Development of an upper limb power assist system using pneumatic actuators for farming lift-up motion. *Journal of System Design and Dynamics*, Vol. 3, No. 5, pp. 781–791, 2009.

[28] E. Yagi, D. Harada, and M. Kobayashi. Upper-limb power-assist control for agriculture load lifting. *International Journal of Automation Technology*, Vol. 3, No. 6, pp. 716–722, 2009.

[29] Y. Imamura, T. Tanaka, Y. Suzuki, K. Takizawa, and M. Yamanaka. Motion-based-design of elastic material for passive assistive device using musculoskeletal model. *Journal of Robotics and Mechatronics*, Vol. 23, No. 6, p. 978, 2011.

[30] T. Tanaka, Y. Satoh, S. Kaneko, Y. Suzuki, N. Sakamoto, and S. Seki. Smart suit: Soft power suit with semi-active assist mechanism-prototype for supporting waist and knee joint. In *Control, Automation and Systems, 2008. ICCAS 2008. International Conference on*, pp. 2002–2005. IEEE, 2008.

[31] Keijiro Y., Kazuhito H., Mineo I., and Takashi M. Development of power assisting suit for assisting nurse labor. *Transactions of the Japan Society of Mechanical Engineers. Series C*, Vol. 67, No. 657, pp. 1499–1506, may 2001.

[32] Takashi H., Takayuki T., and Yumeko I. Smart suit lite integrated with jockey protector and verification of its load reduction. *Hippophile*, No. 47, pp. 27–33, 2012.

[33] X. Zhang and M. Hashimoto. Synchronization based control for walking assist suit-evaluation on synchronization and assist effect. *Key Engineering Materials*, Vol. 464, pp. 115–118, 2011.

[34] Y. Wada, Y. Kaneko, E. Nakano, R. Osu, and M. Kawato. Quantitative examinations for multi joint arm trajectory planning?using a robust calculation algorithm of the minimum commanded torque change trajectory. *Neural Networks*, Vol. 14, No. 4, pp. 381–393, 2001.

[35] *"A human interface for stride control on a wearable robot"*. Proceedings of The 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2009), pp.4067-4072, 2009.

[36] Takahiko M. Development of power assist system with motion estimation using model predictive control. *IEICE Technical Report HIP*, Vol. 111, No. 432, pp. 1–6, feb 2012.

[37] T. Yanagisawa, Hirata M., Y. Saitoh, A. Kato, D. Shibuya, Y. Kamitani, and T. Yoshimine. Neural decoding using gyral and intrasulcal electrocorticograms. *NeuroImage*, Vol. 45, No. 4, pp. 1099–1106, 2009.

[38] Toshio T., Osamu F., Akira O., and Shin K. A training system for emg manipulation of prosthetic arms. *The Journal of the Institute of Electronics, Information and Communication Engineers. D-II*, Vol. 83, No. 10, pp. 2030–2039, oct 2000.

[39] Hirohumi M., Takanori U., and Hirokazu Y. Supplementary system of myoelectric upper limb prosthesis that projects artificial limb motion from motion of healthy limb. *IEICE Technical Report. NC*, Vol. 105, No. 419, pp. 83–88, nov 2005.

[40] A. Murai, K. Yamane, and Y. Nakamura. Characterization of motor skill based on musculoskeletal model. *Engineering in Medicine and Biology Society, Annual International Conference of the IEEE*, pp. 6542–6545, 2009.

[41] H. YOKOI. Technologies of developmental cognitive machines and mutually adaptable prosthetic hand. *Journal of the Japan Society for Precision Engineering*, Vol. 76, No. 1, pp. 24–28, 2010.

[42] A. d'Avella and E. Bizzi. Shared and specific muscle synergies in natural motor behaviors. *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 102, No. 8, pp. 3076–3081, 2005.

[43] B.T. Volpe, M. Ferraro, H.I. Krebs, and N. Hogan. Robotics in the rehabilitation treatment of patients with stroke. *Current atherosclerosis reports*, Vol. 4, No. 4, pp. 270–276, 2002.

[44] S.M. Prasad, H.S. Maniar, N.J. Soper, R.J. Damiano, and M.E. Klingensmith. The effect of robotic assistance on learning curves for basic laparoscopic skills. *The American journal of surgery*, Vol. 183, No. 6, pp. 702–707, 2002.

[45] J.F. Veneman, R. Kruidhof, E.E.G. Hekman, R. Ekkelenkamp, E.H.F. Van Asseldonk, and H. van der Kooij. Design and evaluation of the lopes

exoskeleton robot for interactive gait rehabilitation. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, Vol. 15, No. 3, pp. 379–386, 2007.

[46] A. Duschau-Wicke, T. Brunsch, L. Lunenburger, and R. Riener. Adaptive support for patient-cooperative gait rehabilitation with the lokomat. *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on IEEE*, pp. 2357–2361, 2009.

[47] A. Duschau-Wicke, A. Caprez, and R. Riener. Patient-cooperative control increases active participation of individuals with sci during robot-aided gait training. *Journal of neuroengineering and rehabilitation*, Vol. 7, pp. 43–43, 2010.

[48] M.D. Ziegler, H. Zhong, R.R. Roy, and V.R. Edgerton. Why variability facilitates spinal learning. *Journal of Neuroscience*, Vol. 30, No. 32, pp. 10720–10726, 2010.

[49] N. Mitsunaga, C. Smith, T. Kanda, H. Ishiguro, and N. Hagita. Adapting robot behavior for human–robot interaction. *Robotics, IEEE Transactions on*, Vol. 24, No. 4, pp. 911–916, 2008.

[50] A. Tapus, C. Ţăpuş, and M.J. Matarić. User-robot personality matching and assistive robot behavior adaptation for post-stroke rehabilitation therapy. *Intelligent Service Robotics*, Vol. 1, No. 2, pp. 169–183, 2008.

[51] R. Gockley, A. Bruce, J. Forlizzi, M. Michalowski, A. Mundell, S. Rosenthal, B. Sellner, R. Simmons, K. Snipes, A.C. Schultz, et al. Designing robots for long-term social interaction. In *Intelligent Robots and Systems, 2005. IEEE/RSJ International Conference on*, pp. 1338–1343. IEEE, 2005.

[52] M. Wellner, R. Sigrist, and R. Riener. Virtual competitors influence rowers. *PRESENCE: Teleoperators and Virtual Environments*, Vol. 19, No. 4, pp. 313–330, 2010.

[53] A. Filippeschi, E. Ruffaldi, A. Frisoli, C.A. Avizzano, M. Varlet, L. Marin, J. Lagarde, B. Bardy, and M. Bergamasco. Dynamic models of team rowing

for a virtual environment rowing training system. *International Journal of Virtual Reality*, Vol. 8, No. 4, p. 49, 2009.

[54] Y. Takeda, Y. Tanaka, and T. Tsuji. A virtual training sports system for measuring human hand impedance. In *Advanced Intelligent Mechatronics, 2003. Proceedings. IEEE/ASME International Conference on*, Vol. 2, pp. 914–919. IEEE, 2003.

[55] J. Nirme, A. Duff, and P.F.M.J. Verschure. Adaptive rehabilitation gaming system: On-line individualization of stroke rehabilitation. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pp. 6749–6752. IEEE, 2011.

[56] G.A.P. Caurin, A.A.G. Siqueira, K.O. Andrade, R.C. Joaquim, and H.I. Krebs. Adaptive strategy for multi-user robotic rehabilitation games. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pp. 1395–1398. IEEE, 2011.

[57] J. Furusho and T. Kikuchi. A 3-d rehabilitation system for upper limbs" emul", and a 6-dof rehabilitation system" robotherapist", and other rehabilitation system with high safety. *Rehabilitation Robotics (Edited by Sashi S Kommu)*, pp. 115–136, 2007.

[58] D.J. Reinkensmeyer, J.L. Emken, and S.C. Cramer. Robotics, motor learning, and neurologic recovery. *Annual Review of Biomedical Engineering*, Vol. 6, pp. 497–525, 2004.

[59] D. Feygin, M. Keehner, and R. Tendick. Haptic guidance: Experimental evaluation of a haptic training method for a perceptual motor skill. In *Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2002. HAPTICS 2002. Proceedings. 10th Symposium on*, pp. 40–47. IEEE, 2002.

[60] J.L. Emken and D.J. Reinkensmeyer. Robot-enhanced motor learning: accelerating internal model formation during locomotion by transient dynamic amplification. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, Vol. 13, No. 1, pp. 33–39, 2005.

[61] R.S. Sutton and A.G. Barto. *Reinforcement learning: An introduction*, Vol. 1. Cambridge Univ Press, 1998.

[62] J. Peters and S. Schaal. Reinforcement learning of motor skills with policy gradients. *Neural Networks*, Vol. 21, No. 4, pp. 682–697, 2008.

[63] M. Casadio, P. Morasso, V. Sanguineti, and P. Giannoni. Minimally assistive robot training for proprioception enhancement. *Experimental brain research*, Vol. 194, No. 2, pp. 219–231, 2009.

[64] Richard A Schmidt and Tim Lee. *Motor Control and Learning, 5E.* Human kinetics, 1988.

[65] Rieko Osu, Satomi Hirai, Toshinori Yoshioka, and Mitsuo Kawato. Random presentation enables subjects to adapt to two opposing forces on the hand. *Nature neuroscience*, Vol. 7, No. 2, pp. 111–112, 2004.

[66] Jeffrey F Israel, Donielle D Campbell, Jennifer H Kahn, and T George Hornby. Metabolic costs and muscle activity patterns during robotic-and therapist-assisted treadmill walking in individuals with incomplete spinal cord injury. *Physical therapy*, Vol. 86, No. 11, pp. 1466–1478, 2006.

[67] Eric T Wolbrecht, Vicky Chan, Vu Le, Steven C Cramer, David J Reinkensmeyer, and James E Bobrow. Real-time computer modeling of weakness following stroke optimizes robotic assistance for movement therapy. In *Neural Engineering, 2007. CNE'07. 3rd International IEEE/EMBS Conference on*, pp. 152–158. IEEE, 2007.

[68] Suguru A., Sadao K., and Fumio M. Bettering operation of robots by learning. *Journal of Robotic systems*, Vol. 1, No. 2, pp. 123–140, 1984.

[69] L.M. Crespo and D.J. Reinkensmeyer. Haptic guidance can enhance motor learning of a steering task. *Journal of motor behavior*, Vol. 40, No. 6, pp. 545–557, 2008.

[70] A.G. CHOWDHARY and J.H. CHALLIS. Timing accuracy in human throwing. *Journal of Theoretical Biology*, Vol. 201, No. 4, pp. 219–229, 1999.

[71] C.A. Putnam. Sequential motions of body segments in striking and throwing skills: descriptions and explanations. *Journal of biomechanics*, Vol. 26, pp. 125–135, 1993.

[72] T. Flash and N. Hogan. The coordination of arm movements: an experimentally confirmed mathematical model. *Journal of Neuroscience*, Vol. 5, No. 7, pp. 1688–1703, 1985.

[73] Y. Uno, M. Kawato, and R. Suzuki. Formation and control of optimal trajectory in human multijoint arm movement. *Biological cybernetics*, Vol. 61, No. 2, pp. 89–101, 1989.

[74] Rieko O., Yuji U., Yoshiharu K., and Mitsuo K. Calculating optimal criteria for trajectory planning using observed movement data. *IEICE Technical Report. NC*, Vol. 96, No. 331, pp. 71–78, 1996.

[75] G.P. Mylonas, K.W. Kwok, D.R.C. James, D. Leff, F. Orihuela-Espina, A. Darzi, and G.Z. Yang. Gaze-contingent motor channelling, haptic constraints and associated cognitive demand for robotic mis. *Medical image analysis*, Vol. 16, No. 3, pp. 612–631, 2012.

[76] L.L. Cai, A.J. Fong, C.K. Otoshi, Y. Liang, J.W. Burdick, R.R. Roy, and V.R. Edgerton. Implications of assist-as-needed robotic step training after a complete spinal cord injury on intrinsic strategies of motor learning. *Journal of neuroscience*, Vol. 26, No. 41, pp. 10564–10568, 2006.

[77] C. Obayashi, T. Tamei, A. Imai, and T. Shibata. Comparison of experts and non-experts in throwing darts based on optimization criteria. pp. 2647–2650. IEEE, 2009.

[78] L. Weaver and N. Tao. The optimal reward baseline for gradient-based reinforcement learning. *The Seventeenth Conference on Uncertainty in Artificial Intelligence*, pp. 538–545, 2001.

[79] T. Tamei and T. Shibata. Policy gradient learning of cooperative interaction with a robot using user's biological signals. *Advances in Neuro-Information Processing*, pp. 1029–1037, 2009.

[80] T. Tsuji, P.G. Morasso, K. Goto, and K. Ito. Human hand impedance characteristics during maintained posture. *Biological cybernetics*, Vol. 72, No. 6, pp. 475–485, 1995.

[81] G. Wulf, C. Shea, and R. Lewthwaite. Motor skill learning and performance: a review of influential factors. *Medical education*, Vol. 44, No. 1, pp. 75–84, 2010.

[82] L.L. Cai, A.J. Fong, C.K. Otoshi, Y. Liang, J.W. Burdick, R.R. Roy, and V.R. Edgerton. Implications of assist-as-needed robotic step training after a complete spinal cord injury on intrinsic strategies of motor learning. *Journal of neuroscience*, Vol. 26, No. 41, pp. 10564–10568, 2006.

[83] V. Squeri, A. Basteris, and V. Sanguineti. Adaptive regulation of assistance 'as needed' in robot-assisted motor skill learning and neuro-rehabilitation. In *Rehabilitation Robotics (ICORR), 2011 IEEE International Conference on*, pp. 1–6. IEEE, 2011.

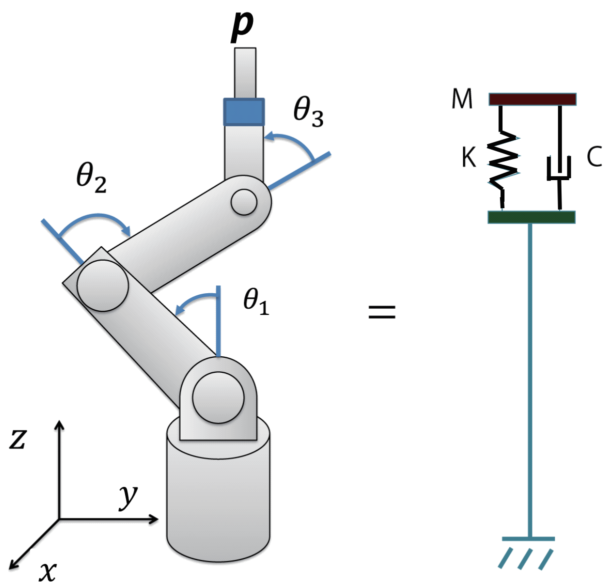



Figure 34. Assistive robot's hand as virtual impedance control

# Appendix

## Virtual Impedance control

The control dynamics of the robot hand is given by

$$\ddot{\boldsymbol{p}} \;=\; \frac{C}{M}\dot{\boldsymbol{p}} + \frac{K}{M}\boldsymbol{p} - \boldsymbol{F}, \tag{13}$$

where $M$, $C$ and $K$ are the mass, viscosity and stiffness set to the robot hand, respectively. In this study, virtual impedance control was applied on only z-axis of the robot hand. Therefore, the velocities of x and y axis were controlled as 0. The controlled joints of the robot were $\boldsymbol{\theta} = [\theta_1, \theta_2, \theta_3]^T$, as shown in Fig.34. The angular velocities sent to the robot, $\theta_1$ and $\theta_2$, were calculated by:

$$\begin{bmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \end{bmatrix} = \boldsymbol{J}(\theta)^{-1}\dot{\boldsymbol{p}}, \tag{14}$$

where $\boldsymbol{J}(\theta)^{-1}$ was the inverse of Jacobian matrix, and $\theta_3$ was obtained by:

$$\theta_3 = -(\theta_1 - \theta_2), \tag{15}$$

so that the robot hand should be kept vertical direction.

# Detail of Experimental Environment

This section trace the history of experimental environment and describes detail of it.

## 1st Season: Inovation Center 3F

Figure 35. Experimental Environment: Inovation Center 3F

Fig. 35 shows the measurement environment for the dart throwing. This area is used following period, -2010 Nov. In the begining, this area is unfitted for behavior experiment. Becuase, It is not covered arround the experiment space by curtain. The control desks are put in experimental space. The subject dees not undisturbedly train by other persons. Therefore, I set up the control desk out of experimental space and put the curtain along the sides of the space.

Figure 36. Experimental Environment: A111 Robot Experiment Room1

## 2nd Season: A111 Robot Experiment Room1

Fig. 36 shows the measurement environment for the experiment of darts throwing. This area is used following period, 2010 Dec. - 2011 April. This area is shared other laboratory. Setting up camera position and calibration were needed in each experiment. After experiment, these systems(Motion capture, DartsBoard, robotic arm) need to put out of this area.

## 3rd Season: B213

Fig. 37 shows the measurement environment for the experiment of darts throwing. This area is used following period, 2011 April - 2011 Dec. This area is set up for the dart throwing learning experiment. late season, the cloasing assistance experiment is began. Therefore, the problem that booking my experiemnt and cloasing assistance experiment is occured. So, the environment for dart throwing experiment need to move the other room like A111.

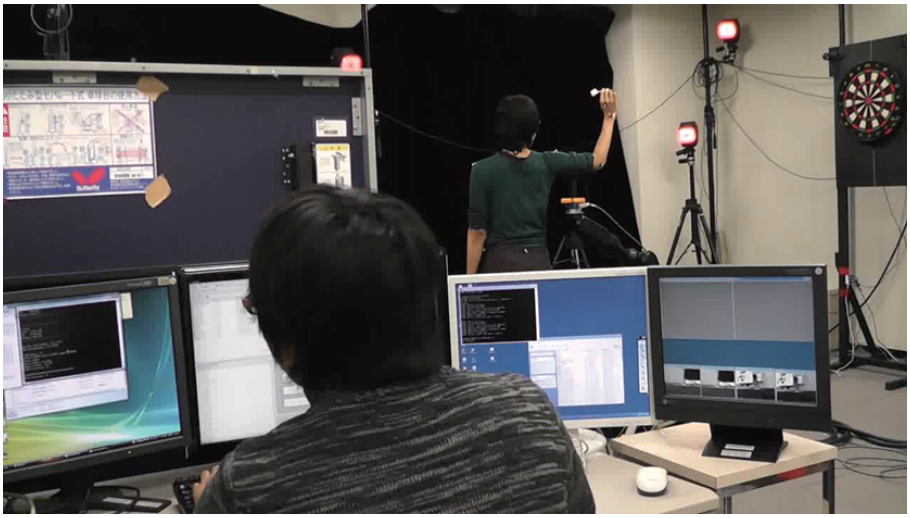


Figure 37. Experimental Environment: B213

Figure 38. Experimental Envirnoment: A111 Robot Experiment Room2

## 4th Season: A111 Robot Experiment Room2

Fig. 38 shows the measurement environment for the experiment. This area is used following period, 2012 April - 2012 June. This area is set up for the dart


Figure 39. Experimental Environement: A111 Robot Experiment Room3

throwing learning experiment.

## 5th Season: A111 Robot Experiment Room3

Fig. 39 and Fig. 40 show the experimental environment and exapmle of the measurement of dart throwing with dart. This area is used following period, 2012 June -. This area is set up to be able to measure the dart and soccer kick motion.

## 6th Season: B213

Fig. 41 shows the measurement environment for the experiment on Sep. 2013. This season we used the B213 again but we need to re-build the measurement environment on Sep. 2013. Becuase the study of the cloasing assistance using the dual-arm robot had used until Aug. 2013. Therefore, dart-measurement system move to the B213 from A111 again and set up the cameras for darts throwing.

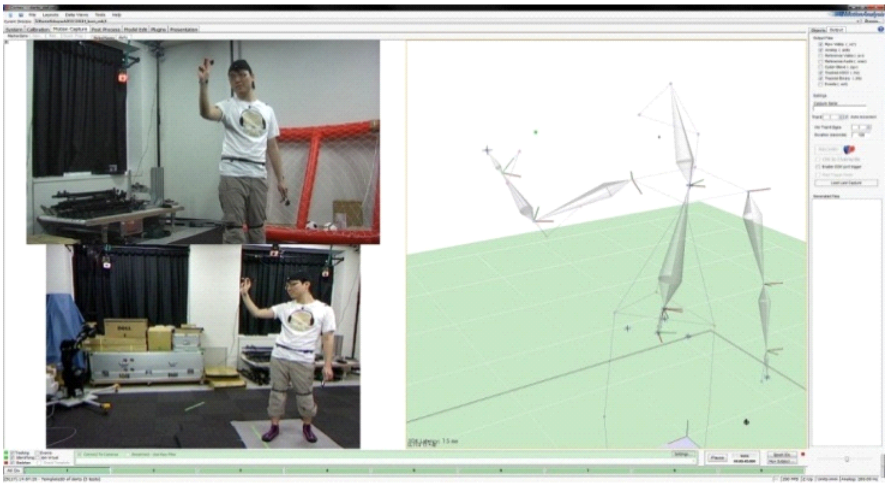


Figure 40. Example of the experiment of measurering of the dart with darts throwing

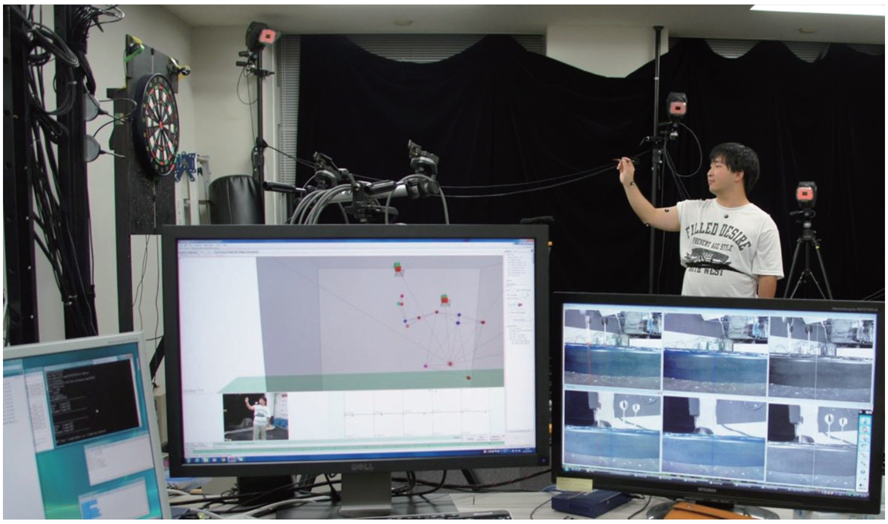


Figure 41. Experimental Environement: A111 Robot Experiment Room3

# Publications

## Journal Papers

- **Chihiro Obayashi**, Tamei Tomoya, Shibata Tomohiro, Assist-as-needed Robotic Training based on Reinforcement Learning - Application to Darts

Throwing, Neural Networks, to appear

## Conference Proceedings (reviewed)

- Tomoya Tamei, **Chihiro Obayashi** and Tomohiro Shibata. Throwing Darts Utilizes the Interaction Torque of the Elbow Joint, Proc. 33rd Annual Int. Conf. IEEE Engineering in Medicine and Biology Society (EMBC 2011), pp.1283-1286, August-Semptember, 2011.

- **Chihiro Obayashi**, Tomoya Tamei, Tomohiro Shibata, Development of an Adaptive Robotic Trainer: Application to Darts Throwing, Proc. 1st Int. Conf. Applied Bionics and Biomechanics (ICABB2010), CD-ROM only, October 2010.

- **Chihiro Obayashi**, Tomoya Tamei, Akira Imai and Tomohiro Shibata, Comparison of experts and beginners in throwing darts based on optimization criteria., Proc. 31st Annual Int. Conf. IEEE Engineering in Medicine and Biology Society (EMBC 2009), pp.2647-2650, September, 2009.

## Conference Proceedings

- **Chihiro Obayashi**，Tomoya Tamei，Tomohiro Shibata, User-Adaptive Robotic Training Accelerates Learning of Darts Throwin, 第23回日本神経回路学会大会 (Neuro 2013)

- **Chihiro Obayashi**，Tomoya Tamei，Tomohiro Shibata, Assist-as-needed Robotic Training based on Reinforcement learning, 第30回日本ロボット学会学術講演会 International Session (RSJ2012).

- 為井智也，**大林千尋**，柴田智広, ダーツ投擲における関節トルク要素の解析, The 21st Annual Conference of the Japanese Neural Network Society.

- **Chihiro Obayashi**，Tomoya Tamei，Tomohiro Shibata, Development of Adaptive Robotic Trainer for Novices of Darts Throwing, 第29回日本ロボット学会学術講演会 International Session (RSJ2011).

- Shibata Tomohiro, **Obayashi Chihiro**, Tamei Tomoya , Adaptive Robotic Training for Darts based on Comparisons of Experts and Non-experts, 第28回日本ロボット学会学術講演会 International Session (RSJ2010).

- **大林千尋**, 為井智也, 柴田智広, 池田和司, ダーツ投擲動作における熟達者と非熟達者の比較. 第24回 生体・生理工学シンポジウム論文集 (BPES2009), pp.295-298.

## 雑誌記事

- 柴田 智広, 船谷 浩之, 中村 彰宏, **大林千尋**, 指さし制御でルンバを自由に制御する「Kinect とフリーなライブラリによるルンバの制御と 2D ゲームの製作」,Interface 2012 年 1 月号 pp.107-111

## 予算獲得

- 健康支援プロジェクト「生駒の歩き方」, **プロジェクトリーダー**, CICP 2008 (Creative and International Competitiveness Project), 創造力と国際競争力を育む情報科学教育コア「プロジェクト型研究提案」

## TA・RA

- TA 計算神経科学 (M2) 2009

- TA 学際領域特論 B　論文作成法・発表法 (M2) 2009

- TA 情報生命学 (D1) 2010

- RA 2010 - 2012