WORKSHOP

OPTIMALITY PRINCIPLES AND ADAPTATION IN HUMANOID ROBOTIC CONTROL

OCTOBER 7, 2012

CONTENTS

ORGANIZERS

MAIN ORGANIZER

Olivier Sigaud
Institut des Systèmes Intelligents et de Robotique
Université Pierre et Marie Curie
Pyramide – Tour 55 - Boîte courrier 173
4 Place Jussieu, 75252 PARIS cedex 05 – France
Email : olivier.sigaud@upmc.fr

CO-ORGANIZERS

Serena Ivaldi (Univ. Pierre et Marie Curie, France)

serena.ivaldi@isir.upmc.fr

Bastien Berret (Italian Institute of Technology, Italy)

bastien.berret@iit.it

Francesco Nori (Italian Institute of Technology, Italy)

francesco.nori@iit.it

ABSTRACT

The availability of new mechatronic platforms such as humanoid robots has driven robotics control research to new paradigms where optimal control of movements and adaptation to perturbations or unforeseen situations are becoming the key challenges. In these new lines of research, inspiration from human motor control and learning capability is playing a central role. But the corresponding frameworks come with very large computational costs so a lot of effort is required in designing practical methods that can address meaningful control and learning problems up to the whole body control of humanoids.

STATEMENT OF OBJECTIVES

The goal of this workshop is to get an accurate picture of the current state of the art and discuss the hot questions of the application of optimal control and adaptation models of human motor control to robotics:

- The inverse optimal control being ill-posed, how can we determine the cost function underlying human movements?
- Do we truly behave optimally in some sense?
- Shall we address whole body control with standard optimal control approaches, or are there some higher level organization principles to be considered?
- What are the next challenges to be addressed?

INTENDED AUDIENCE

We intend to bring together roboticists, optimal control and machine learning experts interested in the use of these methods in robots, as well as computational neurosciences researchers interested in modeling these processes in animals and humans, so as to generate a fruitful discussion between both communities.

LIST OF TOPICS

- Optimal control methods and their derivatives (OFC, SOC, SOFC...)
- Reinforcement learning methods for robotics
- Practical implementations of the above methods in robots
- Humans to humanoids transfer
- Models of human movements based on optimality : successes and limitations

REFERENCES

Ivaldi, S.; Sigaud, O.; Berret, B.; Nori F. (2012) **From Humans to Humanoids: the Optimal Control framework**. Paladyn. Journal of Behavioral Robotics. DOI: 10.2478/s13230-012-0022-3 Pages 1-17.

PROGRAM AT A GLANCE

TIME	SPEAKER
0830- 0845	Olivier Sigaud Introduction
0845- 0900	Stephen Scott
0900- 0915	Putting Sensory back into Voluntary Motor Control
0915- 0930	David Franklin
0930- 0945	Adaptive feedforward and feedback strategies in sensorimotor control
0945- 1000	Etienne Burdet
1000- 1015	A theory of interactive motor behaviors
1015- 1030	POSTER TEASERS
1030-1100	BREAK
1100- 1115	Alexander Terekhov
1115- 1130	Can we learn what the brain optimizes?
1130- 1145	Evangelous Theodorou
1145- 1200	Robustness of optimal control predictions to robotic and neuromuscular control
1200- 1215	Jan Babic
1215-1230	Robot skill synthesis through human sensorimotor learning
1230- 1400	LUNCH
1400- 1415	Katja Mombaur
1415-1430	Identifying optimality principles in human locomotion for humanoid robots
1430- 1445	Jean-Paul Laumond
1445- 1500	The geometric dimensions of action
1500- 1515	Mike Mistry
1515- 1530	Exploiting Redundancy to Optimize the Task Space
1530- 1600	BREAK

1600- 1615	Sethu Vijayakumar
1615- 1630	Optimality in Compliant Actuation: Theory and Practice
1630- 1645	Jan Peters
1645- 1700	Machine Learning of Motor Skills for Robotics
1700- 1715	Martin Butz
1715- 1730	How Do We Interact With Objects: Optimal Control or Weighted Integration of Multiple Biases?
1730-1745	Serena Ivaldi, Bastien Berret, Francesco Nori, Olivier Sigaud
1745-1800	Conclusions

Time schedule:

- morning: 8:30 -12:30 (30min for coffee-break)
- afternoon: 14:00 -18:00 (30 min for coffee-Break)

INVITED TALKS

JEAN-PAUL LAUMOND (LAAS-CNRS, TOULOUSE, FRANCE)

The geometric dimensions of action

Abstract:

By considering first that motions are continuous functions from time to space (i.e. trajectories), and second that actions are compositions of motions, actions appear as sequences of trajectories. The images of the trajectories in spaces are paths. Paths represent geometric traces left by the motions in spaces. The reasoning holds for real space as well as configuration space.

A simple path embodies the entire action. It integrates into a single data structure all the complexity of the action. The decomposition of the action into sub-actions (e.g., walk to, grasp, give) appears as the decomposition of the path into sub-paths. Each elementary sub-path is selected among an infinite number of possibilities within some sub-manifolds. From this perspective the questions are:

- *Motion Segmentation: what are the invariant sub-manifolds that define the structure of a given action?*
- *Motion Generation: among all the solution paths within a given sub-manifold (i.e. among all the possibilities to solve a given sub-task) what is the underlying law that converges to the selection of a particular motion?*

The talk overviews some results obtained in this framework and illustrated from the HRP2-14 humanoid platform.

****

ETIENNE BURDET (IMPERIAL COLLEGE, LONDON, UK)

A theory of interactive motor behaviors

Abstract:

While the motor interaction between a robot and a human, or between humans, has important implications for society as well as promising applications, little research has been devoted to its investigation. In particular, it is important to understand the different ways two agents can interact and be able to generate suitable interactive behaviours. This presentation will thus introduces a framework for the description and implementation of interactive behaviours of two agents performing a joint motor task. A taxonomy of interactive behaviours will be defined, which uses a task classification based on simple questions, and cost functions that represent the way each agent interacts. The role of an agent interacting during a motor task can be directly explained from the cost function this agent is minimising and the task constraints. This framework will be used to interpret and classify previous works on human-robot motor interaction. Its

implementation power will be demonstrated by simulating representative interactions of two humans and comparing with experimental results.

****

JAN BABIC (JOSEF STEFAN INSTITUTE, LJUBLJANA, SLOVENIA)

Robot skill synthesis through human sensorimotor learning

Abstract:

In this talk, I will introduce a concept of obtaining complex robot motions based on the human sensorimotor learning capabilities. The idea is to include the human in the robot control loop and to consider the target robotic platform as a tool that can be iteratively controlled by a human. Provided with an intuitive interface between the human and robot, the human learns to perform a given task using the robot. The skilled control of the robot by the human provides data that are used for construction of an autonomous controller that controls the robot independently of the human. To demonstrate the applicability of the concept, I will present several examples including statically stable reaching and reactive postural control obtained for a humanoid robot. Besides, I will also explain how the interfaces built for the robot skill synthesis can be effectively used in the opposite direction to investigate human motor control mechanisms employed by the central nervous system during the full body motion.

****

DAVID FRANKLIN (UNIVERSITY OF CAMBRIDGE, UK)

Adaptive feedforward and feedback strategies in sensorimotor control

Abstract:

The human sensorimotor control system has exceptional abilities to perform skilful action despite ever changing conditions. I will discuss how this adaptability can result through intrinsic feedback mechanisms in two different ways: sensory feedback driving feedforward adaptation; and feedforward adaptation in turn adapting the feedback responses and tuning them to the environment. When we learn a new skill, sensory feedback from previous errors are incorporated into the feedforward motor command to gradually refine our movements. I will describe a model of motor learning based on the simultaneous optimization of stability, accuracy and efficiency. This model of motor learning offers new insights as to how the brain controls the complex musculoskeletal system and iteratively adjusts motor commands to improve motor skills with practice. However learning can also be used to adjust intrinsic feedback control. I will discuss new experiments in which we examine the control and modulation of involuntary visuomotor responses. These rapid visuomotor responses generate corrective arm movements to

small positional changes in the visual feedback of the hand during movements. I will describe experiments demonstrating that these feedback responses can be precisely tuned by the sensorimotor control system to the environment.

****

ALEXANDER TEREKHOV (ISIR, UPMC, PARIS, FRANCE)

Can we learn what the brain optimizes?

Abstract:

Asking a subject to reach the nose with a fingertip we normally do not expect him or her to pass the arm beneath the knee, though such solution is nonetheless eligible. We find some movements to be convenient and others to be not, and usually when there is something we want to do, we prefer to do that in a convenient way. This suggests that the brain associates a cost with every movement and tends to minimize this cost in motor actions. The question of what is being minimized by the brain is in the air for the last 50 years after it was first raised by Nubar and Contini (1961). Since that time numerous candidate cost functions were proposed for the whole spectrum of motor activities and different levels of analysis, ranging from the activation of the individual motor units within a single muscle to the whole body trajectory control in walking. The problem was analyzed using the formalism of static optimization and optimal control in deterministic and stochastic models. Somehow discouraging seems the fact that for the same motor task there were several cost functions proposed, which could fit the experimental data equally well. In this concern, we ask the question of whether it is at all possible to say for sure what the brain optimizes, at least theoretically. Say, we could measure the activation of every receptor and the stretch of every muscle fiber in as many motor tasks as we needed, would this allow us to determine the cost function(s) used by the brain? This question has fundamental importance for current motor control: the negative answer would put in doubt existing attempts to find out what the brain optimizes, the positive answer would encourage researchers to look for less demanding conditions. The problem of finding a cost function from the experimental data is called inverse optimization. The central question addressed in the presentation will be when, if ever, the inverse optimization problem can have a unique solution. We will formulate the Theorem of Uniqueness for a class of inverse optimization problem. This theorem provides conditions on the hypothetical experimental data, whose satisfaction guarantees that the cost function can be determined unambiguously. The limitations of the theorem will be also discussed in the talk. To make the reasoning more apparent we will consider the inverse optimization problem for the finger force sharing in prehension and finger pressing tasks. We will provide simple examples of how the violation of the conditions of the theorem induces non-uniqueness in solving the inverse optimization problem. We will also report the recent results obtained for grasping. In particular, we will show how the cost function determined from the experimental data for the four fingers grasp can equally well explain the experimental data in the three fingers grasping. It will be concluded that yes, theoretically we can determine the cost function optimized by the brain, but the question of interpreting this cost function is left to the researcher.

****

STEPHEN SCOTT (QUEEN'S UNIVERSITY, KINGSTON, ONTARIO, CANADA)

Putting Sensory back into Voluntary Motor Control

Abstract:

There have been many ways to interpret neural activity in primary motor cortex with the last 25 years largely focused on identifying neural representations (i.e. muscles versus movements). The results highlight myriad representations are present with little consensus on what this means. Recent theories based on optimal control have been influential for interpreting voluntary control and emphasize the importance of online sensory feedback for guiding motor action. My talk will highlight our recent studies that highlight the surprising sophistication of online control in humans including knowledge of limb mechanics, scaling to spatial target location, and avoidance of obstacles in the environment. As well, I will highlight how primary motor cortex provides a key role in this sophisticated use of sensory feedback to guide motor action.

****

EVANGELOS THEODORU (UNIVERSITY OF WASHINGTON, SEATTLE)

Robustness of optimal control predictions to robotic and neuromuscular control

Abstract:

How sensitive optimal predictions are with respect to underlying models of neuromuscular function? How close robotic systems are to neuromuscular systems and what is the role that robotics can play towards the reverse engineered of the brain? Inspired by these questions this talk has two parts: On the theory side I will discuss theoretical developments on iterative path integral control and present Information theoretic interpretations and extensions based on the fundamental relationship between free energy and relative entropy. The aforementioned relationship provides an alternative view of stochastic optimal control theory which does not rely on the Bellman principle. On the application side I will show the applicability of the proposed algorithms to control of biomechanical and robotic systems and try to address the questions of sensitivity and the role of robotics and optimal control towards a better understanding of neuromuscular function.

****

KATJA MOMBAUR (RUPRECHT-KARLS-UNIVERSITÄT HEIDELBERG, GERMANY)

Identifying optimality principles in human locomotion for humanoid robots

Abstract:

A lot of effort in robotics is put into learning complex motor skills from humans. One way to do the transfer between nature and robotics is to use dynamical models and optimization. Inverse optimal control techniques can be used to understand and identify the underlying optimality criteria of human motions based on measurements. The established optimization criteria can then be used to generate optimal motions for the robot - after an appropriate scaling taking into account the different dynamic constraints of the robot - by solving and optimal control problem based on the robot model. Inverse optimal control problems are difficult from a mathematical point of view, since they require to solve a parameter identification problem inside an optimal control problem. We show different examples of human locomotion studies in which we have identified optimality principles using a bilevel optimization approach.

****

MIKE MISTRY (UNIVERSITY OF BIRMINGHAM, UK)

Exploiting Redundancy to Optimize the Task Space

Abstract:

Humans have at their disposal a plethora of degrees of freedom for completing task objectives. Bernstein famously referred to this redundancy as the "degrees-of-freedom problem": how does the nervous system cope with the problem of an indeterminate mapping from goals to actions? Recently, however redundancy has not been viewed as problematic, but rather as a boon for assisting in task achievement. For example, Todorov and Jordan have shown how motion variability is often pushed into redundant space, acting as a "noise buffer" for reducing variability in task space. In this talk I will focus on the relationship between task space and redundancy through the framework of operational space control. Rather than treating redundancy as a merely a passive buffer for handling noise or disturbances, I will discuss how redundancy can be actively controlled to assist in task achievement, as well as to realize certain optimization criteria. For example, I will show how environmental contacts (i.e. constraint forces) can be naturally exploited to reduce actuation effort. For underactuated systems, such as humanoids, task space forces often cannot be applied directly. In this case, I will show how redundant motion is useful for generating forces at passive degrees of freedom and producing task space motion.

****

SETHU VIJAYAKUMAR (UNIVERSITY OF EDINBURGH, UK)

Optimality in Compliant Actuation: Theory and Practice

Abstract:

Variable Impedance modulation allows additional degrees of redundancy in planning and executing dynamic tasks. However, there are interesting machine learning challenges in computing optimal policies to exploit the natural dynamics -- this includes finding the right representation of the dynamics, learning and adapting plant models from data and of course, spatio-temporal optimization of this (expanded) variable impedance policy space. This has to be done in conjunction with complex, non-linear actuator constraints. We will demonstrate our learning and optimization framework and some recent results on its application to dynamic movement tasks on various robotic platforms.

****

JAN PETERS (TECHNISCHE UNIVERSITAET DARMSTADT, GERMANY)

Machine Learning of Motor Skills for Robotics

Abstract:

Autonomous robots that can assist humans in situations of daily life have been a long standing vision of robotics, artificial intelligence, and cognitive sciences. A first step towards this goal is to create robots that can learn tasks triggered by environmental context or higher level instruction. However, learning techniques have yet to live up to this promise as only few methods manage to scale to high-dimensional manipulator or humanoid robots. In this talk, we investigate a general framework suitable for learning motor skills in robotics which is based on the principles behind many analytical robotics approaches. It involves generating a representation of motor skills by parameterized motor primitive policies acting as building blocks of movement generation, and a learned task execution module that transforms these movements into motor commands. We discuss learning on three different levels of abstraction, i.e., learning for accurate control is needed to execute, learning of motor primitives is needed to acquire simple movements, and learning of the task-dependent "hyperparameters" of these motor primitives allows learning complex tasks. We discuss task-appropriate learning approaches for imitation learning, model learning and reinforcement learning for robots with many degrees of freedom. Empirical evaluations on a several robot systems illustrate the effectiveness and applicability to learning control on an anthropomorphic robot arm. These robot motor skills range from toy examples (e.g., paddling a ball, ball-in-a-cup) to playing robot table tennis against a human being.

****

MARTIN BUTZ (UNIVERSITY OF TÜBINGEN, GERMANY)

How Do We Interact With Objects: Optimal Control or Weighted Integration of Multiple Biases?

Abstract:

The traditional control approach postulates that motor control strives for optimality when interacting with the environment. One prominent optimality factor thereby is the end-state comfort effect, that is, interactions focus on ending in a somewhat comfortable final posture. We review several experiments from our group that shows that this optimality criterion, if at all, is only part of the whole story and that a Weighted Integration of Multiple Biases model can indeed account for the data more effectively. This opens a discussion on how, which, and when particular criteria influence behavioral decision making and motor control.

****

POSTERS

- Bruno Damas, Lorenzo Jamone, Jose Santos-Victor and Atsuo Takanishi. "Online Learning of Switching Models Using Multi-valued Function Approximation".
- Yuval Tassa, Tom Erez, Emo Todorov. "Synthesis of Robust Behaviors through online Trajectory Optimization"
- Nima Shafii, Abbas Abdolmaleki, Nuno Lau, Luis Paulo Reis. "A Robust Closed-Loop Gait for Humanoid Clock-Turning"

Online Learning of Switching Models Using Multi-valued Function Approximation

Bruno Damas[1,2], Lorenzo Jamone[3], José Santos-Victor[1] and Atsuo Takanishi[3,4]

I. Introduction

The future of humanoid robots is to become efficient helpers for humans, both in the execution of everyday tasks and in the accomplishment of tedious and dangerous works. Driven by this vision, researchers have been challenged to design more and more complex robots, that show an increasing number of degrees of freedom and sensors [1], [2]; these robots should be able to cope with the unstructured environment in which humans daily live and act. In particular, the ability to reliably control the end-effectors (e.g. hands, feet) in the task space (i.e. the space in which tasks are defined) is fundamental for most practical operations. As control signals are applied to the robot joints, a mapping from the joints space $\mathbf{q}$ to the task space $\mathbf{x}$ is required: the relation $\mathbf{x} = f(\mathbf{q})$ is known as forward kinematic model. In general, this model can be obtained analytically if an accurate description of the system is available. However, building analytical models for current humanoid robots is difficult and time-consuming, especially with robots of increasing complexity. Clearly, discrepancies between the analytical model and the real system affect the performances of the control, arising from the lack of knowledge of certain hard to measure physical parameters (*e.g.* friction) to highly non-linear physical interactions, such as actuator nonlinearities, soft or deformable parts and unmodeled mass distributions [3]; moreover, the properties of the system can change over time, either in a gradual or abrupt way. In such complex situations we must resort to modern supervised learning techniques to provide these systems with the necessary representation capability. In particular, it is desirable that robots are able to: i) estimate this model online from autonomously gathered data and ii) account for both gradual and abrupt changes of the system. The development and implementation of online learning approaches that can support these abilities is a key challenge for the dissemination of humanoid robots in our society.

Several learning approaches have been recently investigated to estimate the robot kinematics or body schema, either focusing on forward [4]–[7] or inverse models [8]. The work of [4] and [5], for instance, proposes learning algorithms for the kinematics of robotic structures using visual information, and in [7] the authors introduce a method based on active learning. Salaun *et al.* [6] propose learning the Jacobian for the forward kinematic function and then inverting it for control in the task space, handling redundancies using null space projections, this way allowing to solve multiple tasks simultaneously.

Some other works focus specifically on the problem of adapting the robot body schema under different tools operation: in [9] a simple 2-joints planar manipulator is controlled using an analytical model of the Jacobian, and when a tool is added to the kinematic chain the corresponding Jacobian is obtained through multiplication of the analytical Jacobian by a linear constant matrix, which is learned exploiting the temporal integration of visual and tactile information during motor exploration. Another approach is proposed in [10], where a recurrent neural network parametrized with the length of the tool is used to estimate the inverse kinematics of a humanoid robot. However, the length of the tool must be known in advance to train and query the neural network. Additionally, training is done using circular trajectories in a fixed plane: this procedure learns a subspace of much lower dimensionality than the joint space dimension being used.

This work investigates how to learn the kinematics of a humanoid robot and use it for task space control, considering, in particular, a reaching task (i.e. positioning the end-effector in the 3D Cartesian space) with tools of different lengths. Switching models, like the forward kinematics of a robotic arm under the use of different tools, can be naturally represented by multi-valued functions, where each function branch represents a different model to learn, and where different output solutions can be obtained from the same query input point. Using such representation has obvious advantages over the single-value model, which expects a unique input-output relation: whenever a model switching takes place (*i.e.*, a change of the end-effector tool), a single-valued learning algorithm has to fully learn the mapping from inputs to outputs again, forgetting the previous learned map. This is very ineffective: it introduces a period of adaptation for learning the new model, even if it has been previously presented to the algorithm. In principle, a multi-valued function approximation algorithm can avoid these issues, as it has no need to forget a branch of the switching model whenever a new one is presented for training.

In this work such kind of multi-valued functions are learned from sensory data using the Infinite Mixture of Linear Experts (IMLE) algorithm [11], a novel online learning algorithm that is particularly suited for these kind of multi-valued

[1]B. Damas and J. Santos-Victor are with the Instituto de Sistemas e Robótica, Instituto Superior Técnico, Lisboa, Portugal. {bdamas,jasv}@isr.ist.utl.pt

[2]B. Damas is with Escola Superior de Tecnologia de Setúbal, Portugal.

[3]L. Jamone and A. Takanishi are with the Faculty of Science and Engineering, Waseda University, Tokyo, Japan lorejam@liralab.it, contact@takanishi.mech.waseda.ac.jp

[4]A. Takanishi is with the Humanoid Robotics Institute, Waseda University, Tokyo, Japan

functions. It is, at its core, a probabilistic algorithm that uses a generalized expectation-maximization (GEM) procedure to update its parameters, fitting an infinite mixture of linear experts to an online stream of training data. Both input space partition and local linear models update are automatically performed through the GEM iterations: some additional priors are also defined to ensure a proper regularization of the mixture parameters, and an outlier probabilistic model is used to decide when to activate a new linear expert in the mixture. Its only assumptions about the training data nature is that it can be approximated by a mixture of local linear models: this naturally allows for multi-valued function learning, as the different branches describing such function can be approximated by different experts sharing the same input region. Obtaining multi-valued predictions is done by a probabilistic clustering procedure, using a statistical test to find the correct number of solutions for a given query. During online operation, a sensible way of picking a single solution from the set of multi-valued predictions provided by the algorithm consists in comparing these solutions to the previous observed training sample and choosing the nearest one: this takes advantage of the typical temporal correlation of the training stream of data that typically occurs in online learning schemes.

To test the hypothesis that multi-valued function approximation can lead to better results under a switching models framework than a single-valued regression algorithm, we chose to compare the IMLE algorithm with the Locally Weighted Projection Regression (LWPR) algorithm [12], a online non-linear function approximation algorithm widely used for robotic learning problems: it's excellent memory requirements and computational complexity have made LWPR a reference online learning algorithm for single-valued functions. Like IMLE, LWPR uses a collection of local linear models to represent the function to be learned, thus avoiding the need to keep all training points in memory, and consequently allowing the use of online learning schemes, particularly suited for real-time operation. It uses a gradient descent on the prediction error, based on a stochastic leave-one-out cross-validation algorithm, to adapt the distance metrics of the receptive fields that partition the input space. Every time a new training point fails to activate at least one of these receptive fields for a given threshold a new linear models is created. Within each receptive field, a linear relation from input to output is obtained via an incremental partial least squares algorithm that efficiently deals with redundant and high-dimensional input spaces. Final predictions are obtained through a weighted average of the linear models individual predictions, where the weights are provided by the corresponding receptive fields activations. This algorithm can only perform single-valued prediction: this is a direct consequence of the optimization criterion for the adaptation of the distance metrics of each receptive field, based on the minimization of the (single-valued) prediction error.

II. Experimental Results

We conducted experiments using the iCub Dynamic Simulator [13] (a snapshot is displayed in figure 1), actuating 7 of its joints (4 DOFs for the right arm, 3 DOFs for the waist) to control the end-effector position, corresponding to either the hand or the tool tip, in the 3D Cartesian space. The robot kinematics is estimated online and used for task space control, relying on the Jacobian pseudo inversion; null-space projection is employed to handle the system redundancy. This solution for learning based control is similar to the one adopted in [6], and the choice for the redundancy resolution was originally proposed in [14]: a main task is realized in the Cartesian space (i.e. positioning of the end-effector) and a secondary task is pursued in joint space (i.e. keeping the joints as far as possible from the physical limits).

At each control step the Jacobian $J(\mathbf{q})$ mapping motor velocities to task velocities is needed to calculate motor velocities that fulfil the desired tasks, and consequently it has to be obtained from the current estimated map relating joint and task spaces (i.e. the forward kinematics), using either the IMLE or the LWPR algorithm in the online training and prediction process. To acquire a model of the forward kinematics the robot first performs motor babbling in joint space. The performances of the task space controller are then evaluated by executing a test movement, consisting of following a given task space trajectory. The test is performed controlling the position of either the hand or the tip of a tool. The right image in figure 1 shows a representative subset of the 3D positions of the end-effector during the motor babbling using different tools (a 28 cm long stick tool and a 48x30 cm L shaped tool): during this process the robot moves to random reference points in the joint space using a low-level joint position control, while acquiring joint values and respective 3D positions of the end-effector and presenting these training pairs to the learning algorithms.

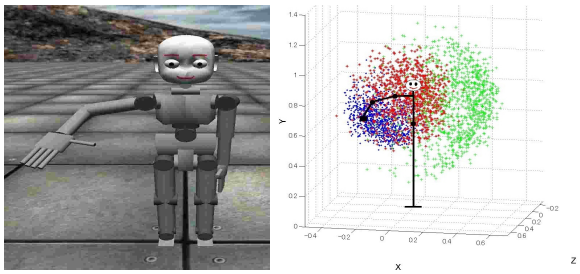


Fig. 1. On the left: the iCub Simulator. On the right: a sketch of the robot kinematics, with the 3D positions of the end-effector during motor babbling (blue dots, using the hand, red stars, using a stick tool, and green crosses, using a L shaped tool, respectively).

While learning the robot kinematics mapping we sometimes change the tool used as the end-effector, without signalling it to the learning algorithms. More precisely, we first perform babbling without any tool for 100,000 training samples. Then we attach the stick tool and learn the mapping for another 100,000 points, and finally we detach the tool and perform random babbling for the remaining training

points. Figure 2 shows the resulting learning curves for both IMLE and LWPR. Multi-valued regression superiority is here

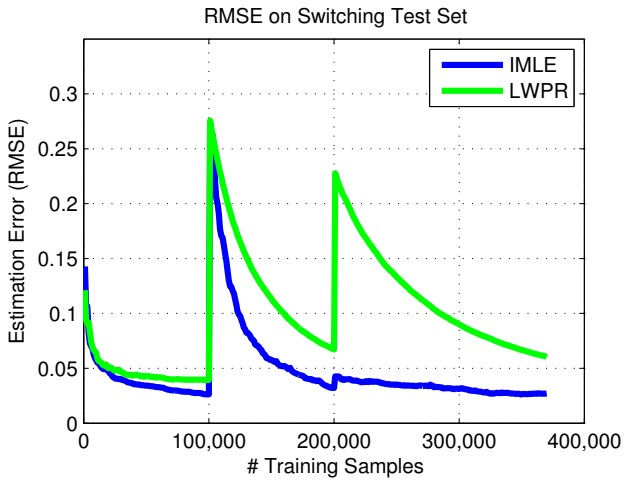


Fig. 2. RMSE on the test set when the tool used during the training is changed (test set is chosen to match the corresponding training set).

evident, specially after detaching the tool: LWPR has to learn the full kinematics mapping all over again. IMLE, on the other hand, does not need to restart learning after the two different situations are presented to it: as can be seen in the figure, after removing the tool (200,000 training points) the RMSE suffers almost no change.

We then alternated random babbling with a) no tool attached; b) stick tool attached and c) L shaped tool attached. We performed babbling in each data set for 100,000 data points, and at the end of each babbling phase we tried to follow a cube-shaped task space trajectory using different end-effector configurations. Figures 3–8 show the resulting trajectories.

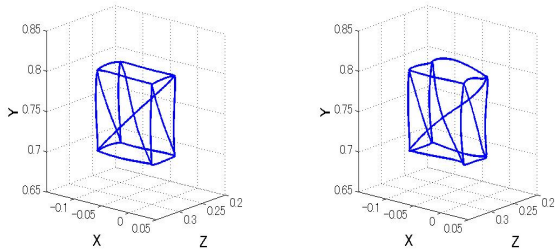


Fig. 3. Task space trajectory of the hand during the test movement, after motor babbling without tool was performed. On the left: using IMLE. On the right: using LWPR.

As can be seen, LWPR can't correctly follow the task space trajectory whenever a different tool configuration than the last babbling configuration is used. IMLE, on the other hand has no trouble at all dealing with such environment, as it has the ability to store the different kinematics relations in its internal model. Note also that LWPR start to degrade its performance after several tool switches take place: this can be seen in Figure 8, where the task space trajectory is

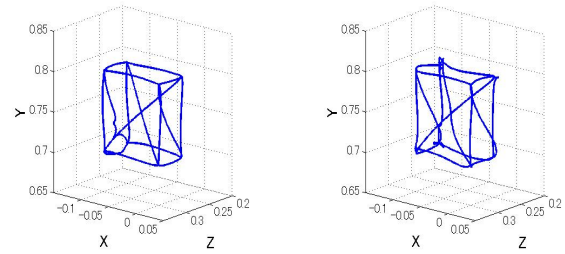


Fig. 4. Task space trajectory of the tip of the stick tool during the test movement, after motor babbling was performed first without and then with the tool. On the left: using IMLE. On the right: using LWPR.

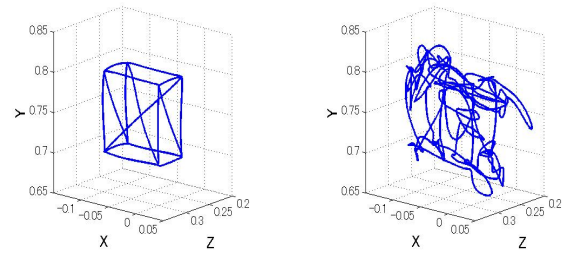


Fig. 5. Task space trajectory of the hand during the test movement, after motor babbling was performed first without and then with the tool. On the left: using IMLE. On the right: using LWPR.

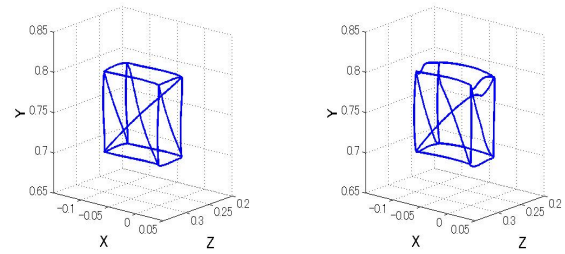


Fig. 6. Task space trajectory of the hand during the test movement, after motor babbling was performed first without tool, then with tool, and then again without tool. On the left: using IMLE. On the right: using LWPR.

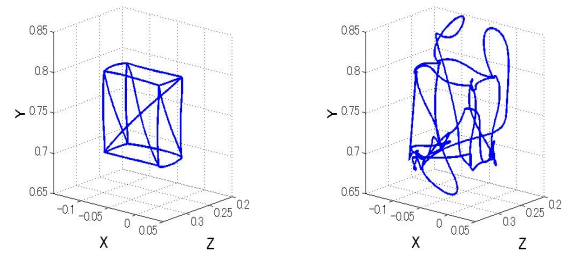


Fig. 7. Task space trajectory of the tip of the stick tool during the test movement, after motor babbling was performed first without tool, then with tool, and then again without tool. On the left: using IMLE. On the right: using LWPR.

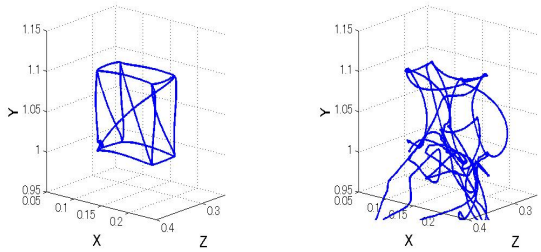


Fig. 8. Task space trajectory of the tip of the L shaped tool during the test movement, after motor babbling was performed first without tool, then with the stick tool, then without tool, and then with the L shaped tool. On the left: using IMLE. On the right: using LWPR.

performed using the same tool as the one used in the most recent training phase.

III. Conclusions

We presented a novel approach to learn the kinematic model of a redundant robot for task space control that can cope with the flexible inclusion of tools of different lengths and shapes. Modelling the forward kinematics as a multi-valued function and using IMLE, an online multi-valued function approximation algorithm, to learn the corresponding switching model, introduced a great improvement over the usual single-valued learning, which requires relearning every time a change in the model occurs, corresponding to a non signalled tool change. Differently from previous works in the literature, no assumptions were made about the kinematic properties of the tool; also, no information was given to the robot about the current tool being used, or when a change or removal of the tool was performed. Simulation results show the effectiveness of the proposed strategy: after some motor experience was acquired through autonomous exploration, the robot could switch from one tool to another without degradation in the control performances, with the possibility of dynamical inclusion of new tools in the learned model. In this work learning is performed both during motor babbling and during the control; however, the motor babbling part can be limited or eliminated, due to the online nature of IMLE, to realize a complete goal-directed exploration, as proposed in [15]. This strategy is very general and can be applied to any robot which needs to be controlled in a specific task space. In the simulation experiments we control the 3D Cartesian position of the robot hand, and the same can be realized on a real robot if an absolute position sensor (i.e. tracker) is available on the hand, or if the hand position is extracted from cameras or laser devices. In alternative, the task space position can be described directly with camera coordinates to realize visually guided reaching in humanoid robots equipped with in-eye cameras. Following this latter approach, we aim at implementing the proposed learning system on real humanoid robots, namely iCub [1] and Kobian [2]: this will be the topic of a future work.

References

[1] G. Sandini, G. Metta, and D. Vernon, "The icub cognitive humanoid robot: An open-system research platform for enactive cognition," in *Lecture Notes in Computer Science*. Springer Berlin / Heidelberg, 2007, vol. 4850, pp. 358–369.

[2] N. Endo, K. Endo, M. Zecca, and A. Takanishi, "Modular design of emotion expression humanoid robot kobian," in *18th CISM-IFToMM Symposium on Robot Design, Dynamics, and Control (ROMANSY)*, Udine, Italy, 2010.

[3] J. Peters and S. Schaal, "Learning Operational Space Control," *Robotics: Science and Systems (RSS 2006)*, 2006.

[4] M. Hersch, E. Sauser, and A. Billard, "Online learning of the body schema," *International Journal of Humanoid Robotics*, vol. 5, no. 2, pp. 161–181, 2008.

[5] J. Sturm, C. Plagemann, and W. Burgard, "Unsupervised body scheme learning through self-perception," in *International Conference on Robotics and Automation*, 2008.

[6] C. Salaun, V. Padois, and O. Sigaud, "Control of reundant robots using learned models: an operational space control approach," in *International Conference on Intelligent Robots and Systems (IROS)*, St. Luis, USA, October, 11-15 2009, pp. 878–885.

[7] R. Martinez-Cantin, M. Lopes, and L. Montesano, "Body schema acquisition through active learning," in *International Conference on Robotics and Automation*, Anchorage, Alaska, USA, 2010.

[8] A. D'Souza, S. Vijayakumar, and S. Schaal, "Learning inverse kinematics," in *International Conference on Intelligent Robots and Systems (IROS)*, Maui, Hawaii, USA, Oct. 29 - Nov. 03 2001, pp. 298–303.

[9] C. Nabeshima, Y. Kuniyoshi, and M. Lungarella, "Adaptive body schema for robotic tool-use," *Advanced Robotics*, no. 20(10), pp. 1105–1126, 2006.

[10] M. Rolf, J. J. Steil, and M. Gienger, "Learning flexible full body kinematics for humanoid tool use," in *Int. Symp. Learning and Adaptive Behavior in Robotic Systems*, 2010.

[11] B. Damas and J. Santos-Victor, "An Online Algorithm for Simultaneously Learning Forward and Inverse Kinematics," in *IEEE/RSJ International Conference on Intelligent Robots and Systems, 2012. IROS 2012, Accepted.*, 2012.

[12] S. Vijayakumar, A. D'Souza, and S. Schaal, "Incremental Online Learning in High Dimensions," *Neural Computation*, vol. 17, no. 12, pp. 2602–2634, 2005.

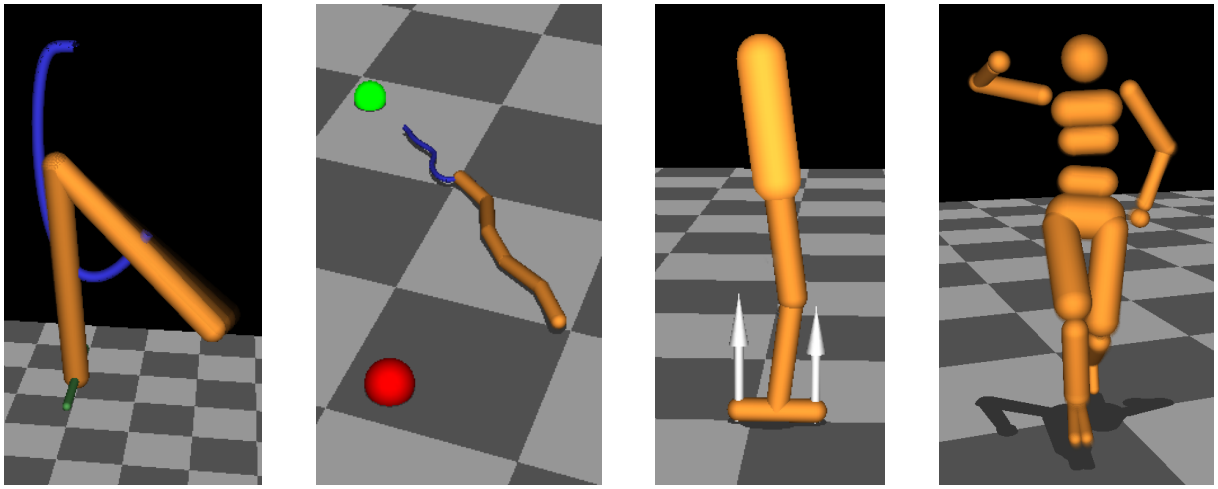
[13] V. Tikhanoff, P. Fitzpatrick, G. Metta, L. Natale, F. Nori, and A. Cangelosi, "An open source simulator for cognitive robotics research: The prototype of the icub humanoid robot simulator," in *Workshop on Performance Metrics for Intelligent Systems*, National Institute of Standards and Technology, Washington DC, August 19-21 2008.

[14] A. Ligeois, "Automatic supervisory control of the configuration and behavior of multibody mechanisms," *Transactions on System, Man and Cybernetics*, no. 7, pp. 868–871, 1977.

[15] L. Jamone, L. Natale, K. Hashimoto, G. Sandini, and A. Takanishi, "Learning task space control through goal directed exploration," in *Robotics and Biomimetics (ROBIO), 2011 IEEE International Conference on*. IEEE, 2011, pp. 702–708.

Synthesis of Robust Behaviors through Online Trajectory Optimization

Yuval Tassa, Tom Erez, Emo Todorov
Computer Science & Engineering
University of Washington, Seattle, USA
{tassa,etom,todorov}@cs.washington.edu



(a) acrobot (b) swimmer (c) hopper (d) humanoid

Fig. 1: Single frames from a movie demonstrating our results: dl.dropbox.com/u/56715/humanoid.m4v

I. Summary

We demonstrate an online trajectory optimization method and software platform applicable to complex humanoid robots performing challenging tasks such as getting up from an arbitrary pose on the ground and recovering from very large disturbances using acrobatic maneuvers. The resulting behaviors, illustrated in the movie referenced in the figure caption, are computed only 7 times slower than real time, on a laptop computer. The movie also shows results on the acrobot problem, planar swimming and one-legged hopping. These simpler problems can already be solved in real time, without precomputing anything and without specifying heuristic approximations to the value function. We demonstrate both robustness to state perturbations and robustness to large modeling errors (unknown to the controller).

II. Motivation

The framework of optimal control makes it possible to specify high-level task goals through simple cost functions and synthesize the details of the behavior and control law automatically. In practice however, optimal control is rarely applied to systems with high-dimensional state spaces – due to the "curse of dimensionality". The curse is particularly problematic for humanoid robots, whose state space is so large that no control scheme (optimal or not) can explore all of it in advance, and prepare suitable responses for every situation. The most impressive results to date have been achieved by local methods, which side-step the curse of dimensionality at the price of local minima and open-loop control. A further complication in the context of locomotion are contact phenomena – which are inherently discontinuous and thus difficult to handle automatically.

Here we describe our recent efforts to overcome the above limitations. Our goal is to construct intelligent feedback controllers that are not limited to the vicinity of precomputed trajectories, and to do so fully automatically, without need for manual specification of contact phenomena or any other aspects of the behavior. We achieve this through online trajectory optimization, also known as model-predictive control (MPC) or receding-horizon control. The idea is to optimize a trajectory up to some time horizon starting at the current state, apply the initial control signal along this trajectory, and repeat. The previous solution is used to warm-start the optimizer, which often yields convergence after a single step, if the minimum has been well-tracked. Contacts are handled automatically by contact smoothing methods we have developed.

The main difficulty in applying MPC is the need to (re)optimize movement trajectories in real time. This may seem impossible for a 3D humanoid performing a complex task. However advances on multiple fronts have brought us surprisingly close to this goal. Controlling the 23-dof humanoid illustrated in the figure (in the task of getting up from an arbitrary pose) is currently only 7 times slower than real time, on a single desktop machine. The computational bottleneck is in the derivatives of the dynamics, computed via finite differencing, which in turn is "embarrassingly parallel". Thus if we were to connect 10 such computers in a cluster, we should be able to control this humanoid in real time. We have already implemented the distributed version of our code, and hope to be able to demonstrate such real time control before the workshop.

III. State of the Art

In domains such as chemical process control where the dynamics are sufficiently slow and smooth – and thus online trajectory optimization is already feasible – MPC is the method of choice [1]. In robotics, however, the typical timescales of the dynamics are orders of magnitude smaller. Furthermore, many robotic tasks involve contact phenomena that present a serious challenge to optimization-based approaches. As a result, MPC is rarely used to control complex robots. Autonomous helicopter flight [2] is a recent example of the power of MPC applied to robotics, although that system is lower-dimensional and smooth. Another illustration is our work on bouncing two ping-pong balls on the same paddle [3]. This task involves contacts but the dimensionality is still relatively low.

The ball-bouncing example relies on a heuristic approximation to the optimal value function, which is used as a final cost applied at the MPC horizon. In general there is a natural tradeoff between how good the value function approximation is, and how much work the MPC machinery has to do [4]. Here we focus on the case when no such approximation is available, and all the work is done by MPC. Thus our results are in some sense worst-case results, and the performance of our method can be improved by using suitable value function approximations. One way to obtain such approximations automatically is to apply machine learning methods to the vast amount of data generated by the MPC controller.

Another approach to generating intelligent feedback (beyond linear) with trajectory optimizers are aggregation methods, like trajectory libraries [5] and LQR trees [6]. While this approach is promising, our guess is that it will eventually run into the curse of dimensionality. Presumably as the volume of state space grows with the dimension, either the number of local controllers will have to grow exponentially, or the region of validity of each local controller will have to grow exponentially, either way susceptible to an explosion in computational complexity.

IV. Our Approach

The results we will demonstrate are enabled by advances on multiple fronts. Our new physics simulator, called MuJoCo, was used to speed up the computation of dynamics derivatives. MuJoCo is a C-based, platform-independent, multi-threaded simulator tailored to control applications. We developed several improvements to the iterative LQG method for trajectory optimization [7] that increase its efficiency and robustness. We also developed several models of contact dynamics [8], [9], [10] which yield different trade-offs between physical realism and speed of simulation/optimization. We introduced cost functions that result in better-behaved energy landscapes and are more amenable to trajectory optimization. Finally, we developed a MATLAB-based GUI where the user can modify the dynamics model, cost function or algorithm parameters, while interacting in real time with the controlled system. We have found that the hands-on familiarity with the various strengths and weaknesses of the MPC machinery is invaluable for proper parameter tuning.

References

[1] M. Diehl, H. Ferreau, and N. Haverbeke, "Efficient numerical methods for nonlinear mpc and moving horizon estimation," *Nonlinear Model Predictive Control*, p. 391, 2009.

[2] P. Abbeel, A. Coates, M. Quigley, and A. Y. Ng, "An application of reinforcement learning to aerobatic helicopter flight," in *Advances in Neural Information Processing Systems 19: Proceedings of the 2006 Conference*, 2007, p. 1.

[3] P. Kulchenko and E. Todorov, "First-exit model predictive control of fast discontinuous dynamics: Application to ball bouncing," in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, 2011, p. 2144–2151.

[4] T. Erez, Y. Tassa, and E. Todorov, "Infinite horizon model predictive control for nonlinear periodic tasks," *Manuscript under review*, 2011.

[5] C. Liu and C. G. Atkeson, "Standing balance control using a trajectory library," in *IEEE/RSJ International Conference on Intelligent Robots and Systems, 2009. IROS 2009*. IEEE, Oct. 2009, pp. 3031–3036.

[6] R. Tedrake, "LQR-Trees: feedback motion planning on sparse randomized trees," in *Proceedings of Robotics: Science and Systems (RSS)*, 2009.

[7] E. Todorov and W. Li, "A generalized iterative LQG method for locally-optimal feedback control of constrained nonlinear stochastic systems," in *Proceedings of the 2005, American Control Conference, 2005.*, Portland, OR, USA, 2005, pp. 300–306.

[8] E. Todorov, "Implicit nonlinear complementarity: a new approach to contact dynamics," in *International Conference on Robotics and Automation*, 2010.

[9] Y. Tassa and E. Todorov, "Stochastic complementarity for local control of discontinuous dynamics," in *Proceedings of Robotics: Science and Systems (RSS)*, 2010.

[10] E. Todorov, "A convex, smooth and invertible contact model for trajectory optimization," in *2011 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, May 2011, pp. 1071–1076.

A Robust Closed-Loop Gait for Humanoid Clock-Turning

Nima Shafii, Abbas Abdolmaleki, Nuno Lau, Luis Paulo Reis

Abstract—Turn-in-place or clock-turning is a basic motion in humanoids robots path planning, however, it can be considered as one of the most complicated tasks in biped locomotion studies. A biped robot needs to move all leg joints in all three - transverse (axial), frontal (lateral) and sagittal -planes. This paper presents a model-free approach based on Fourier series to generate walking trajectories, both, in the foot positional space and the joint angular space. By using trunk stability, our emphasis is on making a robot perform "Clock-Turning" motion more stable and fast. A Genetic algorithm was used to optimize joint trajectories produced by Truncated Fourier Series (TFS) in the simulation environment. Finally, Hill Climbing optimization is obtained on real robot to find locally optimal gait in foot positional trajectories. Multiple experiments were conducted to demonstrate the effectiveness of the proposed method on simulated and real robot.

Keywords: Bipedal Locomotion, Gait Generation, Gait Optimization.

I. Introduction

Humanoid robots are organized in a human-like fashion; they have to deal with dynamic constraints such as balance, as well as geometric constraints. They have larger number of joints (redundancy) which makes them more attractive to researchers. While wheeled robot locomotion is not adapted to many human environments such as, stairs and areas littered by many obstacles, this redundancy enables humanoid robots to avoid obstacles, joint limits, and attain more desirable postures. Therefore, by using biped locomotion, humanoid robots can function and perform their tasks easier than wheeled robot in areas designed for people. In the future, biped robots should have ability to maneuver easily in any environment.

Therefore, besides forward walking, other motions like beside walking and turning should come into research interests of biped locomotion researchers. Biped robots to reach a destination need to change their direction to avoid obstacles or following an specified path in most scenarios. Kuffner et al [1] proposed an online footstep planning method for biped robots to move and avoid obstacles in environment. This work mostly is focused on obstacle avoidance for robot where the balance and locomotion is not the matter.

S. N. is with the Artificial Intelligence and Computer Science Laboratory, Faculty of Engineering of the University of Porto(FEUP), Porto, Portugal. (e-mail: nima.shafii@fe.up.pt).

L. P. R. is with the School of Engineering, University of Minho, Guimaraes, Portugal. (e-mail: lpreis@fe.up.pt).

A.A. is with the Instituto de Engenharia Electronica e Telematica de Aveiro (IEETA) ,Universidade de Aveiro, Aveiro, Portugal (e-mail: abbas@ua.pt).

L. N. is with the Instituto de Engenharia Electronica e Telematica de Aveiro (IEETA) ,Universidade de Aveiro, Aveiro, Portugal (e-mail: nuno.lau@ua.pt).

Kajita et al. [2] proposed a method to control humanoid walking motion. Their method was based on a simplified 3D inverted pendulum model (LIPM), which overlooked substantial dynamics of robots and has constraints for planning to move the center of mass, e.g. it can't model to move the Center of Mass (CoM) in Z axis. He also presented a turning gait planning, in which the objective of turning is to follow a curve with certain radius [3]. This is called forward-turning or curving. The aim of clock-turning is to change robot's direction at a fixed position. These two kinds of turning are suitable in different situations, due to the fact that in forward-turning, there are no switches between different motion gaits during the whole obstacle avoidance. In this case, robots require more prediction and consideration of accurate position and shape of obstacles. But, clock-turning is able to combine turning and walking to adapt to different shapes of obstacles. Therefore, clock-turning has lower requirements for the humanoid sensor system. Although the clock-turning has been used as a basic motion for humanoid global path planning, there is only one previous work on it [4]. In 2007 Tang et al. tried to investigate turn motion by using linear inverted pendulum model , and preview control of the zero momentum point (ZMP) indicators and inversed kinematics[4]. Like other 3D LIPM approaches, one of the main issues, is that they cannot produce running and jumping movements since it models the support leg as the massles telescopic leg.

In our previous studies [5][6][7], we employed Fourier series and evolutionary algorithms like genetic and PSO algorithm to generate angular trajectory for bipedal direct walking and we verified our method through simulation and experimental results. This method has three significant advantages. First of all, it can be used for every humanoid robot without any mathematical modeling, so in our previous work we called it a model free approach. Second this approach has a low computational complexity and can be implemented on real humanoid robots, which the response of controller should be real time. Finally Fourier series can parameterize a signal well; therefore robot locomotion controller designer can manipulate signals simply by changing parameters of Fourier series in proper times. In addition, the last advantage leads utilizing optimization algorithms to achieve the best parameters of Fourier series to generate proper control signals for locomotion gate planning.

In this paper, a new model-free approach for turning-in-place is proposed. In this method, Fourier series are used to model joints angular trajectories on the simulated robot. Genetic algorithm is used to optimize the parameter of the model on the simulated robot.

Learned trajectories of simulation results are reconfigured

again in order to apply on the real robot. Fourier series is used to remodel the movement of the knee for support leg of a real robot in foot positional space. This can lead us to produce the movement of the CoM in *Z* axis. A sensory feedback based on the High frequency response to the torso accelerometer is also implemented, in order to have the trunk stability control on the sagittal plane. Hill climbing as a local search technique is applied to optimize parameters of the model on a real robot.

II. Turn Angular Trajectories

As it was mentioned, First we test our method on a simulated humanoid robot and then simulation results are transferred to a real robot. The robot model in this study is the NAO robot and the simulation is performed by Rcssserver3d [8]. The robot model has 22 DOFs with a height of about 57cm, and a mass of 4.5kg. Schematic view of our humanoid robot is shown in fig. 1. DOFs 2,3, and 4 move on Sagital plane and DOFs 1 and 5 move in Frontal plane while DOF 6 moves on transverse plane.

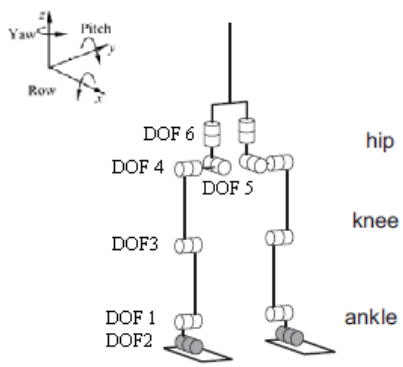


Fig. 1. Schematic view of the lower body of the humanoid robot

Human turn motion can be investigated in many aspects; turning trajectory used here is divided into two types: Positional trajectory and angular trajectory. Foot positional trajectory and Joints Angular trajectory. Angular trajectories provide the angle of each joint at each time slice. In order to achieve the shape of angular trajectory of a turning motion, we start by analyzing angular trajectories results from previous works based on the 3D LIPM approach, which was explained in the introduction and can be found in [4]. Biped angular trajectories of two joints; hip and knee in sagittal, frontal and transverse plane from a humanoid turning are shown in Fig 2,3 and 4 respectively.

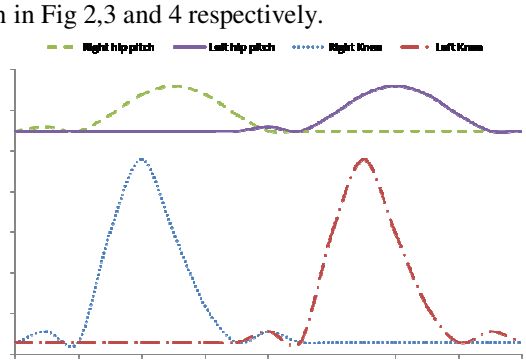


Fig. 2. Hip angle trajectories in sagittal,frontal and transerve planes

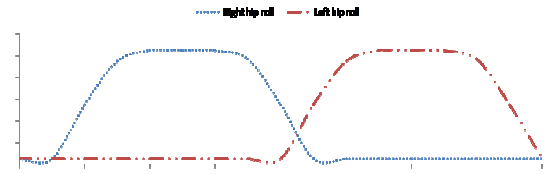


Fig. 3. Knee angle trajectories in sagittal plane

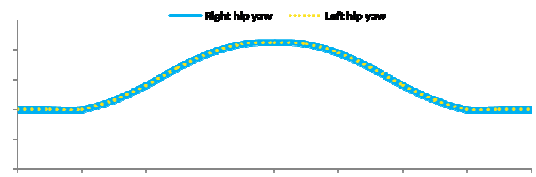


Fig .4 Hip angle trajectory in transverse plane

Since a gait is a cyclic, periodic motion of the joints of a legged robot, it requires the sequencing and coordination of the legs to obtain reliable locomotion. In other words, gait is the temporal and spatial relationship between all the moving parts of a legged robot [13]. Therefore, all gaits angular trajectories are periodic and Fourier Series can be used to model and generate the mentioned trajectories.

In the following sections, we explain how to generate references trajectories for these DOFs in mentioned planes separately to achieve a stable turn-in-place motion for humanoid robot.1

A. *Movements in Sagittal Plane*

As mentioned in previous sections, there are three DOFs in each leg which move on sagittal plane; one in the hip, one in the ankle and one in the knee. The model for generating angular trajectories in sagittal plane is very similar to Truncated Fourier series (TFS) model which was presented in 2009 [5]. In that work, similar to [7], foot in sagittal plane was kept parallel to the ground by using ankle joint. This is done in order to avoid collision. Therefore, ankle trajectory can be calculated by hip and knee trajectories and ankle DOF parameters are eliminated.

In that model, and also according to Fig.2 and Fig. 3. each signal has an offset. In addition, the trajectories for both legs are identical in shape but are shifted in time relative to each other by half of the turning period. So by producing trajectory of one leg the other leg's trajectory can be calculated. The Fourier Series (FS) for generating hip and knee trajectories in sagittal plane are formulated as below "(1),"

$$\theta_{hx} = A_{hx} \,.\, sin(w_{hx}\, t) + C_{hx}\,,\; W_{hx} = \frac{2\pi}{T_{hx}}$$

$$If(\theta_{hx} < C_{hx})$$

$$\theta_{hx} = C_{hx} \qquad (1)$$

$$\theta_{k} = A_{k} \,.\, sin(w_{k}\, t) + C_{k}\,,\; W_{k} = W_{hx}$$

$$If(\theta_{k} < C_{k})$$

$$\theta_{k} = C_{k}$$

In these equations, A_{hx} and A_k are constant coefficients for generating signals. The h_x and k index stands for hip and knee in sagittal plane respectively. Also C_{hx} and C_k are signal offsets and T_{hx} is assumed the period of hip trajectory. As it is mentioned in [4], all joints except "hip yaw" joint have equal movement frequency and the frequency of Fourier Series (FS) for "hip yaw" joint is half of other joints. Therefore, the $W_{hx}= W_k= 2\pi/T_{hx}$ equation can be derived. Fourier series parameters in sagittal plane are A_{hx}, A_k , C_{hx}, C_k and W_{hx}..

B. Movements in Frontal plane

Each leg has two DOFs which locomote on Frontal plane; in hip and ankle. the hip roll angular trajectory in a single turning period was illustrated in Fig.2. Similar to sagittal plane, foot stays parallel to the ground. To do so, angle of ankle in frontal plane stays equal to hip´s angle of the opposite leg. So the trajectory of ankle in frontal plane can be derived from the trajectory of hip angle in frontal plane. Shafii et.al also modeled a similar movement in frontal plane by the TFS [6]. The Fourier Series formula for generating hip trajectories in frontal plane is as follows "(2)".

$$\theta_{hy} = A_{hy} \,.\, sin(w_{hy}\, t) \quad , W_{hy} = W_{hx}=\frac{2\pi}{T_{hx}}$$
$$If(\theta_{hy} < 0) \qquad (2)$$
$$\theta_{hy} = 0$$

The above formula, h_y index representing hip in frontal plane. A_{hy} and W_{hy} are amplitude and frequency of signal respectively. T_{hy} is period of hip in frontal plane. period of hip signal in frontal plane and sagittal plane are equal. W_{hy} is eliminated from our unknown parameter set. Correct value of A_{hy} parameter must be found.

C. Movements in transverse plane

The hip DOF allows moves on the tranverse plane. As has been shown in Fig.4, for both legs the signal for this joint is the same, so we need to determine the parameters of one leg. The NAO robot also has on actuator in order to move in transverse plane. The Fourier Series equations that generate hip trajectory in the transverse plane is formulated below "(3)".

$$\theta_{hr} = A_{hr} \,.\, sin(w_{hy}\, t)+ C_{hr} \quad W_{hr}=W_{hx}/2=\frac{\pi}{T_{hx}}$$
$$If(\theta_{hr} < C_{hr}) \qquad (3)$$
$$\theta_{hr} = C_{hr}$$

In these equations, A_{hr} is constant coefficient for generating signal. The hr index stands for hip in transverse plane. Also C_{hr} is signal offset. As it is mentioned, the movement frequency of FS for hip yaw joint is half of other joints, so the $W_{hr}= W_{hx}/2$ equation can be concluded. Therefore, W_{hr} can be achieved from W_{hx}, and It is eliminated from parameters to be found . Consequently, A_{hr} and C_{hr} should be determined. So parameters of Fourier series to generate proper angular trajectories for turn-in-place motion are $A_{hx}, A_{hy}, A_{hr}, A_k, C_{hx}, C_{hr}, C_k, W_{hx}$ And an optimization algorithms must optimize the 8 dimensions problem to find the best turn gate generator in this stage.

D. Genetic Algorithms For Optimizing Trajectories On the simulated robot

Genetic Algorithms (GA) is a stochastic searching procedure based on the mechanics of natural selection and genetics [9]. GA is used to find the best parameters to generate angular trajectories for robot turning motion. To follow these angular trajectories all individual robot joints are controlled by using proportional derivative (PD) controllers.

Truncated Fourier Series (TFS) has 8 parameters to generate all joints angular trajectories, thus each chromosome has 8 Genes, In Designing of Chromosomes gene's type is considered as double format. Population for each generation is assumed to be 100.

Equation 4 shows fitness function formulation, where the robot is initialized in x=y=0 (0,0) aligned to the horizontal axis where *rotdegree=0* and time duration for turning is 60 *s*.

```
Fitness = rotdegree - 90*dist
```

In the above equation the *rot degree* is the amount of turning at radian and *dist* is the amount of movement from initial place. The GA optimization has been configured with a scattered function for the cross over operator. For mutation, a uniform function has been with mutation rate is assumed 0.06.

Selection method is roulette wheel and reproduction rate is assumed as 0.8.

7 hours after starting GA on a Pentium IV 3 GHz machine with 2 GB of physical memory, the number of generations was 28 and the robot could turn in place 360 degree in 4s with average body speed of around 90 degree/s. In below figures angular trajectory generated by TFS after learning process are shown.

III. REMODEL LEARNED TRAJECTORIES IN ORDER TO APPLY ON THE REAL ROBOT

Gait optimization, using a Genetic algorithm, led the simulated robot to learn how to turn in place. Every Physical simulator contains some simplifications in its real world model; therefore the results of simulation and reality are usually not the same. However, current simulators are quite precise and simulation results have many similarities with those obtained in the real world. This issue is called the reality gap. The reality gap is smaller if the behavior is executed at slow speeds; also the stability of most behaviors is enhanced when using lower speeds.

According to [10] smaller gait with lower amplitude has lower speed and acceleration than bigger gait with higher amplitude. In order to reduce the speed of turn motions, all angular trajectories of the legs (learned by simulation) can be multiplied by a variable assumed as K. The value of this variable can range from zero to one. Zero value produces the lowest speed, the robot will be stopped, and one produces the same trajectories determined by the simulation. The K value can be found by any local search techniques.

The specification of TFS approach, by defining each joint

trajectory independently from the others, does not provide a good model for controlling the foot trajectory. To achieve a more controllable model, the TFS approach specification is converted to use the Leg interface.

Leg interface was presented in 2006. The behavior uses a leg interface to control the leg movements. The leg interface allows specification of the leg positions by using three components: leg angle, leg extension and foot angle. [11]

Leg extension γ is assumed to be the distance between the hip and the ankle. It can be normalized in the range of -1 and 0, $-1 \leq \gamma \leq 0$, where $\gamma = -1$ denotes that the leg is fully extended and $\gamma = 0$ denotes the leg is shortened.

With this new model it is easy to control foot height trajectory. changing the leg extension value of the support foot, robot can produce the movement of the CoM in Z axis, which the classical model of the LIPM can generate this type of the movements. This type of the movement can produce a little jumping which diminishes foot collisions with the ground and results more stable turn behavior. The Fourier Series equations that leg extension trajectory is formulated below "(4)".

$$\gamma = A_{ex} \cdot \sin(w_{ex}\, t) + C_{ex} \quad W_{ex} = W_{hx} \qquad (4)$$

The ex index stands for extension, A_{ex} is constant coefficient for generating signal. Also C_{er} is signal offset. Like the previous equations , the movement frequency of FS is equal to other behaviors, so the $W_{ex=}\ W_{hx}$.
A_{ex} and C_{ex} are the two parameters of the model of the Leg extension generator which the value of them must be found.

A. Posture Control feedback

In this study a simple posture control feedback is also added to make the behavior more robust against external disturbances, such as uneven ground and collisions with obstacles. The feedback is calculated using a filtered value of the accelerometer in the x direction (front); it is based on the rUNSWift Team Report 2010[12]. Equation (5) shows how to calculate the feedback values.

$$filAccelX = A * filAccelX + (A - 1) * accelX * B$$

A and B are parameters of the filtering. $filAccelX$ is summed to two hip pitch joints, to balance the torso. To do so, when the robot is falling to the front, the accelerometer will have a positive value that is added to the hip pitch joints, making the legs to move forward, to compensate. The inverse is also true, when the robot is falling backwards.

B. Hill Climbing For Optimizing parameters 0n the real robot

The model for changing learned trajectories has 5 parameters. The proper value of these 5 parameters is found by using hill climbing optimization. As an instance, turn in place on the real NAO was achieved with K value of 0.3. NAO robot could turn in place 360 degree in 8s stably and average body speed was around 45 degree/s. "Fig. 5" shows clock-turning obtained from GA search in simulation and after adaptation on the NAO Robot while performing clock-turning behavior.

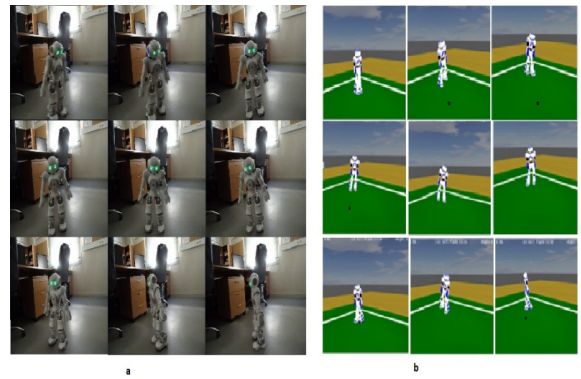


Fig.5. A) NAO Robot was adapted to do clock-turning by using simulation results in lower amplitude: B) Simulated NAO robot follows the learned nominal trajectory to do clock-turning

REFERENCES

[1] J. Kuffner, S. Kagami, K. Nishiwaki, M. Inaba, and H. Inoue, "Online footstep planning for humanoid robots," in *2003 IEEE International Conference on Robotics and Automation Cat No03CH37422*, 2003, no. Icra, pp. 932-937.

[2] S. Kajita, F. Kanehiro, K. Kaneko, K. Yokoi, and H. Hirukawa, "The 3D linear inverted pendulum mode: a simple modeling for a biped walking pattern generation," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2001, pp. 239-246.

[3] S. Kajita, F. Kanehiro, K. Kaneko, K. Fujiwara, K. Yokoi, and H. Hirukawa, "A realtime pattern generator for biped walking," in *Proceedings 2002 IEEE International Conference on Robotics and Automation Cat No02CH37292*, 2002, vol. 1, no. May, pp. 31-37.

[4] Z. Tang, Z. Sun, H. Liu, and M. J. Er, "Clock-turning gait synthesis for humanoid robots," *Journal of Control Theory and Applications*, vol. 5, no. 1, pp. 23-27, Feb. 2007.

[5] N. Shafii,S. Aslani, S.Shiry "Evolution of biped walking using truncated fourier series and particle swarm optimization," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 5949 LNAI, 2010, pp. 344-354.

[6] N. Shafii, L. P. Reis, and N. Lau, "Biped Walking using Coronal and Sagittal Movements based on Truncated Fourier Series," in *RoboCup 2010: Robot Soccer World Cup XIV*, Volume 655., Javier Ruiz-del-Solar, E. Chown, and Paul-Gerhard, Eds. 2011, pp. 324-335.

[7] N. Shafii, Javadi M H S, "A truncated fourier series with genetic algorithm for the control of biped locomotion," in *IEEE/ASME International Conference on Advanced Intelligent Mechatronics, AIM*, 2009, pp. 1781-1785.

[8] J. Boedecker and M. Asada, "SimSpark – Concepts and Application in the RoboCup 3D Soccer Simulation League," *Autonomous Robots*, pp. 174-181, 2008.

[9] M. Shrivastava, "Trajectory Generation Using GA for an 8 DOF Biped Robot with Deformation at the Sole of the Foot," *Journal of Intelligent and Robotic Systems*, vol. 47, no. 4, pp. 197-84, 2007.

[10] N. Shafii, A. Khorsandian, A. Abdolmaleki, and B. Jozi, "An optimized gait generator based on fourier series towards fast and robust biped locomotion involving arms swing," in *IEEE International Conference on Automation and Logistics*, 2009, pp. 2018-2023.

[11] S. Behnke, "Online trajectory generation for omnidirectional biped walking," in *IEEE International Conference on Robotics and Automation, 2006. ICRA 2006.*, 2006, pp. 1597-1603.

[12] A. Ratter, B. Hengst, B. Hall, B. White, and B. Vance, "runswift team report 2010 robocup standard platform league," 2010.