

On Motion Control and Machine Learning for Robotic Assembly

Martin Karlsson



LUND
UNIVERSITY

Department of Automatic Control

Lic. Tech. Thesis TFRT-3274
ISSN 0280-5316

Department of Automatic Control
Lund University
Box 118
SE-221 00 LUND
Sweden

© 2017 by Martin Karlsson. All rights reserved.
Printed in Sweden by Media-Tryck.
Lund 2017

Abstract

Industrial robots typically require very structured and predictable working environments, and explicit programming, in order to perform well. Therefore, expensive and time-consuming engineering work is a major obstruction when mediating tasks to robots. This thesis presents methods that decrease the amount of engineering work required for robot programming, and increase the ability of robots to handle unforeseen events. This has two main benefits: Firstly, the programming can be done faster, and secondly, it becomes accessible to users without engineering experience. Even though these methods could be used for various types of robot applications, this thesis is focused on robotic assembly tasks.

Two main topics are explored: In the first part, we consider adjustment of robot trajectories generated by dynamical movement primitives (DMPs). The framework of DMPs as robot trajectory generators has been widely used in robotics research, because of their convergence properties and emphasis on easy modification. For instance, time scale and goal state can be adjusted by one parameter each, commonly without further considerations. In this thesis, the DMP framework is extended with a method that allows a robot operator to adjust DMPs by demonstration, without any traditional computer programming or other engineering work required. Given a generated trajectory with a faulty last part, the operator can use lead-through programming to demonstrate a corrective trajectory. A modified DMP is formed, based on the first part of the faulty trajectory and the last part of the corrective one. Further, a method for handling perturbations during execution of DMPs on robots is considered. Two-degree-of-freedom control is used together with temporal coupling, to achieve practically realizable reference trajectory tracking and perturbation recovery. In the second part of the thesis, a method that enables robots to learn to recognize contact force/torque transients acting on the end-effector, without using a force/torque sensor, is presented. A recurrent neural network (RNN) is used for transient detection, with robot joint torques as input. A machine learning approach to determine the parameters of the RNN is presented.

Each of the methods presented in this thesis is implemented in a real-time application and verified experimentally on a robot.

Acknowledgments

I would like to thank my supervisor Prof. Rolf Johansson, for your invaluable advice and support throughout this work. Because of your balance between pointing out interesting research directions and giving me freedom to define my work, I look forward to many more days as a PhD student. Thank you for sharing so much of your knowledge and experience.

My co-supervisor Prof. Anders Robertsson, thank you for all your guidance. Three years ago, you patiently taught me how to communicate with the internal controller of an industrial robot. Since then, you have supported me not only with theoretical insights, but also in overcoming practical difficulties, easily encountered in a laboratory environment.

Fredrik Bagge Carlson, I am very lucky to have you as a close colleague. It has been both fruitful and fun to do research together with you. We have a common interest not only in our work, but also in solid-state physics in general and defects in semiconductors in particular, and this has resulted in many interesting discussions and good times. Thank you!

Dr. Björn Olofsson, you deserve deep gratitude for contributing with your energy, your friendliness, and your patience with practicalities in the lab. Thank you for the many times you came to the rescue, when I was lost in the world of Terminal commands.

I would like to acknowledge my colleagues in the RobotLab, Dr. Mahdi Ghazaei Ardakani, Dr. Maj Stenmark, Asst. Prof. Mathias Haage, Prof. Jacek Malec, Assoc. Prof. Elin Anna Topp, Assoc. Prof. Klas Nilsson, Dr. Anders Nilsson, Anders Blomdell and Pontus Andersson, as well as former colleagues, Dr. Magnus Linderoth, Dr. Olof Sörnmo, Martin Holmstrand, Dr. Karl Berntorp and Dr. Andreas Stolt. Thank you for great cooperation.

Dr. Mårten Wadenbäck, thank you for being a fantastic colleague, co-author, and friend, and for introducing me to the world of visual odometry. It was worth all the difficulties with cables and camera settings, and even two minor electric shocks.

I have the honor and pleasure of sharing office with Olof Troeng, Irene Zorzan and Victor Millnert. Thank you for creating such a nice working environment! Thanks to you, I enjoy workdays just as much as weekends.

I am grateful to the administrative staff at the department, Ingrid Nilsson, Mika Nishimura, Monika Rasmusson, Cecilia Edelborg Christensen, and Eva Westin, for your help with various things related to my work.

Leif Andersson, without your deep knowledge in \LaTeX , this thesis, as well as my other publications, would have looked significantly worse. Every time I think that I need to add a new package, you convince me to instead remove several unnecessary ones, always with successful results. On a related note, I thank Dr. Björn Olofsson, Axel Karlsson, Fredrik Bagge Carlson, and Dr. Mahdi Ghazaei Ardakani, for proofreading of this thesis.

I also thank all colleagues at the department. Each one creates a positive, inspiring, and fun working environment, and contributes to interesting discussions about control theory and related topics.

Last but not least, Daniel, Ann-Eli, Axel, and Johan Karlsson, have been a fantastic family through my entire life. You constantly provide me with insights and ideas from the by no means negligible world outside my research field. Heartfelt thanks for all the love and support!

Financial Support

Financial support is gratefully acknowledged from the European Commission, under the Framework Programme Horizon 2020 – within grant agreement No 644938 – SARAFun, and under the 7th Framework Programme – within grant agreement No 606156 – FlexiFab, as well as from the Swedish Foundation for Strategic Research through the SSF project ENGROSS. The author is a member of the LCCC Linnaeus Center, supported by the Swedish Research Council, and the ELLIIT Excellence Center, supported by the Swedish Government.

Contents

1. Introduction	9
1.1 Thesis Outline	10
1.2 Dynamical Movement Primitives	10
1.3 Supervised Machine Learning	11
1.4 Problem Formulation	13
1.5 Thesis Contributions	13
2. Publications	14
3. Discussion and Future Work	17
4. Conclusion	19
Bibliography	20
Paper I. Autonomous Interpretation of Demonstrations...	23
1 Introduction	24
2 Problem Formulation	25
3 Motivating Examples	26
4 Description of the Framework	26
5 Experiments	32
6 Results	34
7 Discussion	34
8 Conclusion	36
References	37
Paper II. Two-Degree-of-Freedom Control for... DMPs	41
1 Introduction	42
2 Preliminaries	43
3 Problem Formulation	45
4 Method	47
5 Simulations	48
6 Implementation of Real-Time Application	49
7 Experimental Setup	51
8 Experimental Results	52

Contents

9	Discussion	53
10	Conclusion	58
	References	58
Paper III. Detection of Contact Force Transients...		61
1	Introduction	62
2	Method	63
3	Experiments	68
4	Results	68
5	Discussion	69
6	Conclusion	72
	References	73

1

Introduction

Compared to humans, typical industrial robots are very good at performing sequences of pre-defined movements, with high speed and high accuracy in the position domain. This has promoted automation of repetitive tasks where position control suffices, such as spray painting and welding. However, standard industrial robots perform well only in carefully structured workcells, specifically designed to fit the robot and the given task. In general, it is required that the task is highly repetitive, and possible deviations from the original plan must have been foreseen by the robot programmer. See, *e.g.*, [Spong et al., 2006; Siciliano et al., 2010] for an introduction to robot modeling and control in general. Even under favorable conditions, traditional robot programming is time consuming and requires expert knowledge. As a result, human labor is still more cost effective than automation for many tasks, such as most assembly tasks. Even though these tasks might appear monotonous and predictable, small tolerances and tiny variations between similar parts make it inadequate to just perform a series of accurate movements. Robotic assembly has been addressed in, *e.g.*, [Björkelund et al., 2011]. Further, there is a trend toward manufacturing a given product in a smaller volume and for a shorter time, and then changing to a new one. These circumstances have motivated the following two research objectives:

1. Enable easier and faster robot programming;
2. Enable robots to take proper action with respect to their surroundings.

In this thesis, research toward these objectives is presented. Paper I mainly addresses Objective 1. More specifically, it is investigated how a human could correct a faulty movement performed by a robot, by demonstrating a desired behavior. Paper II deals with online replanning of robot movements to handle unforeseen events, and hence Objective 2 is mainly addressed. In both Paper I and Paper II, robot motion is modeled by dynamical movement primitives (DMPs). The concept of DMPs is introduced in Sec. 1.2.

It should be noted that the two objectives are partly overlapping. For instance, if a robot is able to replan with respect to its workspace, the robot programmer does

not have to take all eventualities into account, which reduces the required programming work. Vice versa, intuitive means of robot programming could allow for a human to mediate suitable behavior, given certain states or events, to the robot. In Paper III, it is investigated how robots could learn to detect force transients from previous experience, and use this as a decision basis. A closely related topic is force estimation and control, see, *e.g.*, [Olsson et al., 2002; García et al., 2006; Stolt et al., 2012]. In Paper III, explicit programming of the detection model is eschewed, to make the human–robot interaction as easy as possible for the human. Instead, a machine learning approach is used. How machine learning can promote robot programming is described in Sec. 1.3.

1.1 Thesis Outline

This thesis consists mainly of three papers, and it is organized as follows. In Chapter 2, the publications authored or co-authored by the thesis author are listed. A discussion and ideas for continuation of the work are presented in Chapter 3, and a conclusion is presented in Chapter 4.

The first part of the thesis consists of Paper I and Paper II, where the DMP concept is augmented to support corrective demonstrations and enhance replanning capabilities. In this part, demonstrations and robot motion control are mainly considered. The second part of the thesis consists of Paper III, and in contrast to the first part, the main focus is not on motion control. Instead, the aim is to enable recognition of sensor data sequences. Despite this difference between the two parts, all three papers present research toward faster and more intuitive mediation of skills from humans to robots, and consider assembly scenarios in particular.

1.2 Dynamical Movement Primitives

Representation and execution of movements is an important area within robotics. An industrial robot program commonly consists of a sequence of movement instructions, each containing some details that specify the movement, such as velocity and end point. Further, the ability to handle deviations from the planned movement is usually very low. Instead, some motion supervision algorithm would typically stop the robot if it would be too far from its position reference or experience too large joint torques, for instance due to some unexpected physical contact.

To enhance real-time motion modulation, DMPs have been proposed in [Ijspeert et al., 2002; Schaal et al., 2003]. The concept has been inspired by the biological movement models presented in [Giszter et al., 1993; Mussa-Ivaldi, 1999]. It was used for robotic learning from unstructured demonstrations in [Niekum et al., 2015], and for object handover in [Prada et al., 2014]. Trajectory-based reinforcement learning has been applied to automatically tune DMP parameters in, *e.g.*, [Pastor et al., 2013; Kober et al., 2008; Abu-Dakka et al., 2015; Kroemer et al., 2010]. A

DMP is a movement model, defined by a weakly nonlinear dynamical system with attractor behavior, so that the state converges to a desired end point. Although there are other alternatives, the most ubiquitous movement model in the DMP framework, which is the one that Paper I and Paper II proceed from, is based on the following damped-spring system.

$$\tau^2 \ddot{y} = \alpha(\beta(g - y) - \tau \dot{y}) + f(x) \quad (1.1)$$

Here, y denotes robot position, g is the goal position, and α and β are positive constants chosen such that the system is critically damped for $f(x) = 0$. Further, $f(x)$ is a learnable forcing term, with significant magnitude only in a finite time window, that allows for detours before reaching g . The evolution rate is scalable through the time parameter τ , and explicit time dependence is avoided with the phase variable x . Once determined, a DMP can be used as a robot motion controller, by sending control signals so that the robot moves according to the evolution of y . As explained in [Ijspeert et al., 2013], coupling terms are easily incorporated in the DMP framework while retaining its convergence properties, which facilitates online motion modulation with respect to the surroundings of the robot.

Comparison with alternative movement representations

Potential fields and splines are often brought up as two alternatives to DMPs for movement representation. Similar to DMPs, potential fields represent attractor landscapes, with convergence to a goal position and without explicit time dependence. Potential fields have been considered for robot control by many researchers, see, *e.g.*, [Khatib, 1986; Koditschek, 1987; Li and Horowitz, 1999]. Vector fields define the movement based on given positions, but determining the vector field given a desired behavior is not straight forward. Design of potential fields for some obstacle avoidance scenarios has been done in, *e.g.*, [Koditschek, 1987]. Further, while DMPs allow for different control signals from the same position, this can not be achieved with conventional potential fields.

For imitation learning, splines have been widely used. Splines are functions that are defined piecewise by polynomials, and retain smoothness where the polynomials connect. It has been shown in, *e.g.*, [Miyamoto et al., 1996; Wada and Kawato, 2004] that demonstrated trajectories can be represented and successfully reproduced by means of splines. However, online replanning is not supported, and temporal and spatial scaling can be done only by recomputation of the spline polynomials.

1.3 Supervised Machine Learning

Traditionally, computers and robots have been programmed by writing explicit code, specifying sets of rules and behaviors in detail. This works well in predictable scenarios, but most of the tasks that humans perform in their everyday life, are far

too complex to mediate in such fashion. For instance, consider the task of distinguishing whether a certain image represents a car or a bicycle, which is in general easy for humans. Indeed, both categories could take different forms, and hand-crafting the rules for classification from raw image data would not be feasible.

It is better to address such problems with machine learning approaches. This field consists of two major parts; supervised and unsupervised learning, see [Bishop, 2007; Murphy, 2012]. In this thesis, supervised learning is considered. In general, supervised machine learning is used to approximate a given function, $y(x)$, with a parameterized function, $\hat{y}(x|\theta)$, which takes some input data x , and maps it to an output \hat{y} . Here, θ denotes the model parameters. In the example of image recognition, x could be pixel values, y would be the true image category, and $\hat{y}(x)$ could be interpreted as the probability distribution over the two categories, *i.e.*, car and bicycle, given x .

In order to learn $\hat{y}(x)$, the model is exposed to a large data set of examples, called training data. In supervised learning, the training data consist of both input data and the corresponding known outputs, usually manually labeled. In the training phase, the elements of θ are adjusted to fit the training data by means of optimization. A loss function, L , in which some measurement of the error of $\hat{y}(x)$ compared to $y(x)$ is included, is minimized with respect to the model parameters.

Since the training data can only include a small subset of all possible data points, an important aspect of machine learning is generalization, *i.e.*, to predict the output given input not used during training. In order to achieve this, the complexity of the model is typically restricted, by keeping the number of parameters low, or by penalizing the complexity by including it in L . Further, test data, not directly used to optimize the model parameters, are used to estimate how well models generalize. It is, however, common to determine some model hyperparameters based on the performance on test data. Therefore, it is good practice to use yet another data set to investigate the generalizability, without affecting the model in any way. Such data are called validation data.

This general approach is adopted in Paper III, where the aim is to take a step toward more intuitive human–robot interaction. Ideally, a non-expert operator should be able to provide a robot with data, enabling it to learn from experience. Similar to the image classification example, a model is trained to determine the class of the data given x . In particular, x consists of robot joint torques, and the task is to determine whether a certain force/torque transient, acting on the robot end-effector, is present or not. One important difference from the image recognition example is that x consists of a time-series rather than a static representation, which should be taken into account when choosing the structure of the model. In Paper III, a recurrent neural network (RNN) is used as classification model. Prior to the training phase, data are gathered by letting the robot experience the force/torque transient, while logging the joint torques. After that, the data sets are formed by labeling the data. Even though this implies some work by the operator, the required time and traditional programming skills can be reduced significantly with this method compared to ex-

plicitly programmed conditions for classification. Related approaches have been presented in [Rodriguez et al., 2010; Rojas et al., 2012; Stolt et al., 2015], but then, the force/torque acting on the end-effector has been measured directly, which requires a force/torque sensor.

1.4 Problem Formulation

The first aim of this thesis is to answer the question of whether it is possible to automatically interpret a correction, made by an operator, of the last part of a robot trajectory generated by a DMP, while retaining the first part. The human–robot interaction must be intuitive, and the result of a correction predictable enough for its purpose. The result should be a new DMP, of which the first part behaves qualitatively as the first part of the original DMP, whereas the last part behaves according to the corrective demonstration. Discontinuities between the original and corrective trajectories must be mitigated.

Further, it should be investigated whether perturbations of trajectories generated by DMPs could be recovered from, while using control signals of moderate magnitudes only. In the absence of significant perturbations, the behavior should resemble that of the original DMP framework described in [Ijspeert et al., 2013].

The above problems will be addressed by augmenting the original DMP framework. Meanwhile, the benefits of the DMP framework, *i.e.*, scalability in time and space as well as guaranteed convergence to the goal g , should be preserved.

It should also be investigated whether robot joint torques could be used to recognize contact force transients during robotic assembly, despite uncertainties introduced by, *e.g.*, joint friction. Finally, it is desirable to explore how the performance of the detection algorithm is affected by the length of the joint torque sequences used as input.

1.5 Thesis Contributions

The main contributions of this thesis are:

- A framework for modification of DMPs by means of corrective demonstrations;
- An augmentation of the DMP framework that enables recovery from perturbations during DMP execution;
- A machine learning procedure for detecting force/torque transients acting on a robot end-effector, by measuring the robot joint torques.

2

Publications

This licentiate thesis is based on the following three papers.

Paper I

Karlsson, M., A. Robertsson, and R. Johansson (2017). “Autonomous interpretation of demonstrations for modification of dynamical movement primitives”. In: *IEEE International Conference on Robotics and Automation (ICRA)*. May 29–June 3, Singapore.

In this publication, M. Karlsson formulated the method for updating a partly faulty trajectory representation, based on a corrective demonstration. Further, M. Karlsson implemented the method and verified it experimentally. A. Robertsson and R. Johansson contributed with comments on the research and the manuscript.

Paper II

Karlsson, M., F. Bagge Carlson, A. Robertsson, and R. Johansson (2017). “Two-degree-of-freedom control for trajectory tracking and perturbation recovery during execution of dynamical movement primitives”. In: *20th IFAC World Congress*. July 9–14, Toulouse, France. Accepted for publication.

M. Karlsson and F. Bagge Carlson identified the necessity of augmenting the existing DMP framework, to make related research approaches on DMP perturbation recovery practically realizable. M. Karlsson formulated the augmentation, and verified it in simulations and experimentally, while frequently discussing the work with F. Bagge Carlson. Further, F. Bagge Carlson implemented the method presented as an open-source Julia package, which can be found on [Bagge Carlson, 2016]. Example code in Matlab, written by M. Karlsson, can be found on [Karlsson, 2017]. Throughout the work, A. Robertsson and R. Johansson supervised the research and assisted in structuring the manuscript.

Paper III

Karlsson, M., A. Robertsson, and R. Johansson (2017). “Detection of contact force transients during robotic assembly without a force sensor”. Manuscript prepared for submission to review for publication.

M. Karlsson formulated the transient detection method, implemented it, and performed the experimental work. A. Robertsson and R. Johansson contributed with insights regarding related work, provided comments on the research, and assisted in structuring the manuscript.

The following publications, authored or co-authored by the author of this thesis, cover topics in robotics, nonlinear state estimation, and positioning. They are, however, not included in this thesis.

Bagge Carlson, F., M. Karlsson, A. Robertsson, and R. Johansson (2016). “Particle filter framework for 6D seam tracking under large external forces using 2D laser sensors”. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. October 9–14, Daejeon, South Korea.

Haage, M., S. Profanter, I. Kessler, A. Perzylo, N. Somani, O. Sörnmo, M. Karlsson, S. G. Robertz, K. Nilsson, L. Resch, et al. (2016). “On cognitive robot woodworking in SMErobotics”. In: *ISR 2016: 47th International Symposium on Robotics*. VDE. June 21–22, Munich, Germany, pp. 1–7.

Karlsson, F., M. Karlsson, B. Bernhardsson, F. Tufvesson, and M. Persson (2015). “Sensor fused indoor positioning using dual band wifi signal measurements”. In: *European Control Conference (ECC)*. July 15–17, Linz, Austria, pp. 1669–1672.

Karlsson, M., F. Bagge Carlson, J. De Backer, M. Holmstrand, A. Robertsson, and R. Johansson (2016a). “Robotic seam tracking for friction stir welding under large contact forces”. In: *7th Swedish Production Symposium (SPS)*. October 25–27, Lund, Sweden.

Karlsson, M., F. Bagge Carlson, J. De Backer, M. Holmstrand, A. Robertsson, R. Johansson, L. Quintino, and E. Assuncao (2016b). “Robotic friction stir welding, challenges and solutions”. *Welding in the World, The International Journal of Materials Joining*. ISSN: 0043-2288. Submitted to review for publication.

Karlsson, M. and F. Karlsson (2016). “Cooperative indoor positioning by exchange of bluetooth signals and state estimates between users”. In: *European Control Conference (ECC)*. June 29–July 1, Aalborg, Denmark, pp. 1440–1444.

Wadenbäck, M., M. Karlsson, A. Heyden, A. Robertsson, and R. Johansson (2017). “Visual odometry from two point correspondences and initial automatic camera tilt calibration”. In: *12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, Volume 6*. VISIGRAPP. February 27–March 1, Porto, Portugal, pp. 340–346.

3

Discussion and Future Work

Much more work is still required, both in terms of research and engineering, before robot programming could be considered non-problematic. In the future, one natural way to extend the DMP functionality presented here, would be to incorporate trajectory-based learning of suitable actions based on sensor data. Such learning could be warm-started by defining DMP parameters from one initial demonstration. Under the DMP execution, the robot could then deviate from the demonstrated trajectory, based on sensor feedback. Further, albeit the control algorithm in Paper II worked satisfactorily in the simulations and experiments, and some stability properties were addressed, it remains to construct a formal proof of convergence to the goal state.

It would be reasonable to expect that the method for force/torque transient recognition in Paper III would work well for other tasks than that presented, since no assumptions were made in the method design regarding task, parts to be assembled, *etc.*, except that a transient would be generated. However, it remains to evaluate the approach experimentally on various tasks, *e.g.*, assembly of other parts, to verify robustness and generalizability. Further, the data acquisition and labeling is still a bottleneck in this approach, requiring half a working day for the thesis author per task to learn. A natural continuation would therefore be to automatize the work flow as much as possible. For instance, the operator could be aided by a GUI, showing plots of the robot joint torques, and asking the operator to indicate where the transients occur. Perhaps, the operator could be provided with suggestions, already after labeling a few examples, by forming a preliminary, less complex, detection model.

The movement before and after the transient, *i.e.*, moving down in the scenario in Paper III, does not necessarily take much time to implement, but requires traditional coding with the current setup. It would be more accessible, if it could be demonstrated by, *e.g.*, lead-through programming. However, physical contact with the robot arm would affect the joint torques, and by that corrupting the training data and test data. A possible solution for this has been presented in [Ghazaei Ar-

dakani, 2016], where one of the two robot arms has been used for lead-through teleoperation of the other. That approach would allow one robot arm to perform the task, without physical contact with the operator, according to the operator's demonstration on the other arm. Ideally, from such demonstrations, the robot should not only learn recognition of transients, but also desired actions before and after these. It would therefore be valuable to integrate the approach in [Ghazaei Ardakani, 2016] with that in Paper III.

4

Conclusion

The aim of the research presented in this thesis was to facilitate robot programming by enhancing the ability of robots to learn from demonstrations and from experience. More specifically, the DMP framework described in [Ijspeert et al., 2013] was augmented with two new functionalities. First, an algorithm was developed that allowed an operator to correct the last part of a faulty trajectory generated by a DMP, while retaining the first part. The correction could be done in an intuitive way by demonstrating a corrective trajectory. The first part of the resulting DMP behaved as the first part of the original DMP, and the last part behaved according to the correction. Discontinuities between the first and last parts were eschewed by formulating and solving a convex optimization problem. Secondly, a control algorithm that enabled trajectory tracking and perturbation recovery during execution of DMPs was presented. The algorithm was based on a combination of two-degree-of-freedom control and temporal coupling. Since the required control signals were of moderate magnitude only, the controller was practically realizable, which was the main benefit compared to state of the art. In the absence of perturbations, the controller behaved like the original DMP framework.

Further, a machine learning procedure was presented for model-based detection of force/torque transients acting on the robot end-effector without direct force/torque measurement. Instead, robot joint torque sequences were used as model inputs. Therefore, a force/torque sensor was not required, which was the main benefit as compared to previous research. A systematic approach for choosing a suitable length of the input sequences was presented. An RNN was used as classification model for the detection. The approach presented seems promising, since the resulting model showed high performance for the test data as well as during the experiments.

Bibliography

- Abu-Dakka, F. J., B. Nemeč, J. A. Jørgensen, T. R. Savarimuthu, N. Krüger, and A. Ude (2015). “Adaptation of manipulation skills in physical contact with the environment to reference force profiles”. *Autonomous Robots* **39**:2, pp. 199–217.
- Bagge Carlson, F. (2016). *DynamicMovementPrimitives.jl*. Department of Automatic Control, Lund University. URL: <https://github.com/baggepinnen/DynamicMovementPrimitives.jl> (visited on 2017-03-21).
- Bishop, C. (2007). *Pattern Recognition and Machine Learning*. Springer, New York.
- Björkelund, A., L. Edström, M. Haage, J. Malec, K. Nilsson, P. Nugues, S. G. Robertz, D. Störkle, A. Blomdell, R. Johansson, M. Linderoth, A. Nilsson, A. Robertsson, A. Stolt, and H. Bruyninckx (2011). “On the integration of skilled robot motions for productivity in manufacturing”. In: *IEEE/CIRP International Symposium on Assembly and Manufacturing (ISAM)*. May 25–27, Tampere, Finland.
- García, J. G., A. Robertsson, J. G. Ortega, and R. Johansson (2006). “Generalized contact force estimator for a robot manipulator”. In: *IEEE International Conference on Robotics and Automation (ICRA)*. May 15–19, Orlando, FL, pp. 4019–4024.
- Ghazaei Ardakani, M. (2016). *On Trajectory Generation for Robots*. PhD thesis. TFRT–1116–SE, Dept. Automatic Control, Lund University, Lund, Sweden.
- Giszter, S. F., F. A. Mussa-Ivaldi, and E. Bizzi (1993). “Convergent force fields organized in the frog’s spinal cord”. *Journal of Neuroscience* **13**:2, pp. 467–491.
- Ijspeert, A. J., J. Nakanishi, H. Hoffmann, P. Pastor, and S. Schaal (2013). “Dynamical movement primitives: learning attractor models for motor behaviors”. *Neural Computation* **25**:2, pp. 328–373.

- Ijspeert, A. J., J. Nakanishi, and S. Schaal (2002). “Movement imitation with non-linear dynamical systems in humanoid robots”. In: *IEEE International Conference on Robotics and Automation (ICRA)*. Vol. 2. May 11–15, Washington, DC, pp. 1398–1403.
- Karlsson, M. (2017). *DMP perturbation, simulation example*. Department of Automatic Control, Lund University. URL: https://gitlab.control.lth.se/cont-mkr/dmp_perturbation_sim_example (visited on 2017-03-21).
- Khatib, O. (1986). “Real-time obstacle avoidance for manipulators and mobile robots”. *The International Journal of Robotics Research* **5**:1, pp. 90–98.
- Kober, J., B. Mohler, and J. Peters (2008). “Learning perceptual coupling for motor primitives”. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. September 22–26, Nice, France, pp. 834–839.
- Koditschek, D. (1987). “Exact robot navigation by means of potential functions: some topological considerations”. In: *IEEE International Conference on Robotics and Automation (ICRA)*. Vol. 4. March 31–April 3, Raleigh, NC, pp. 1–6.
- Kroemer, O., R. Detry, J. Piater, and P. J (2010). “Combining active learning and reactive control for robot grasping”. *Robotics and Autonomous Systems (RAS)* **58**:9, pp. 1105–1116.
- Li, P. Y. and R. Horowitz (1999). “Passive velocity field control of mechanical manipulators”. *IEEE Transactions on Robotics and Automation* **15**:4, pp. 751–763.
- Miyamoto, H., S. Schaal, F. Gandolfo, H. Gomi, Y. Koike, R. Osu, E. Nakano, Y. Wada, and M. Kawato (1996). “A Kendama learning robot based on bi-directional theory”. *Neural Networks* **9**:8, pp. 1281–1302.
- Murphy, K. P. (2012). *Machine Learning: A Probabilistic Perspective*. MIT press, Cambridge, MA.
- Mussa-Ivaldi, F. A. (1999). “Modular features of motor control and learning”. *Current Opinion in Neurobiology* **9**:6, pp. 713–717.
- Niekum, S., S. Osentoski, G. Konidaris, S. Chitta, B. Marthi, and A. G. Barto (2015). “Learning grounded finite-state representations from unstructured demonstrations”. *The International Journal of Robotics Research* **34**:2, pp. 131–157.
- Olsson, T., J. Bengtsson, R. Johansson, and H. Malm (2002). “Force control and visual servoing using planar surface identification”. In: *IEEE International Conference on Robotics and Automation (ICRA)*. May 11–15, Washington, DC, pp. 4211–4216.
- Pastor, P., M. Kalakrishnan, F. Meier, F. Stulp, J. Buchli, E. Theodorou, and S. Schaal (2013). “From dynamic movement primitives to associative skill memories”. *Robotics and Autonomous Systems* **61**:4, pp. 351–361.

- Prada, M., A. Remazeilles, A. Koene, and S. Endo (2014). “Implementation and experimental validation of dynamic movement primitives for object handover”. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. September 14–18, Chicago, Illinois, pp. 2146–2153.
- Rodriguez, A., D. Bourne, M. Mason, G. F. Rossano, and J. Wang (2010). “Failure detection in assembly: force signature analysis”. In: *IEEE Conference on Automation Science and Engineering (CASE)*. August 21–24, Toronto, Canada, pp. 210–215.
- Rojas, J., K. Harada, H. Onda, N. Yamanobe, E. Yoshida, K. Nagata, and Y. Kawai (2012). “A relative-change-based hierarchical taxonomy for cantilever-snap assembly verification”. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. October 7–12, Vilamoura, Portugal, pp. 356–363.
- Schaal, S., A. Ijspeert, and A. Billard (2003). “Computational approaches to motor learning by imitation”. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **358**:1431, pp. 537–547.
- Siciliano, B., L. Sciavicco, L. Villani, and G. Oriolo (2010). *Robotics: Modelling, Planning and Control*. Springer Verlag, London, UK.
- Spong, M. W., S. Hutchinson, and M. Vidyasagar (2006). *Robot Modeling and Control*. John Wiley & Sons, Hoboken, NJ.
- Stolt, A., M. Linderoth, A. Robertsson, and R. Johansson (2012). “Force controlled robotic assembly without a force sensor”. In: *IEEE International Conference on Robotics and Automation (ICRA)*. May 14–18, St. Paul, MN, pp. 1538–1543.
- Stolt, A., M. Linderoth, A. Robertsson, and R. Johansson (2015). “Detection of contact force transients in robotic assembly”. In: *IEEE International Conference on Robotics and Automation (ICRA)*. May 26–30, Seattle, WA, pp. 962–968.
- Wada, Y. and M. Kawato (2004). “A via-point time optimization algorithm for complex sequential trajectory formation”. *Neural Networks* **17**:3, pp. 353–364.

Paper I

Autonomous Interpretation of Demonstrations for Modification of Dynamical Movement Primitives

Martin Karlsson Anders Robertsson Rolf Johansson

Abstract

The concept of dynamical movement primitives (DMPs) has become popular for modeling of motion, commonly applied to robots. This paper presents a framework that allows a robot operator to adjust DMPs in an intuitive way. Given a generated trajectory with a faulty last part, the operator can use lead-through programming to demonstrate a corrective trajectory. A modified DMP is formed, based on the first part of the faulty trajectory and the last part of the corrective one. A real-time application is presented and verified experimentally.

Originally published in the 2017 IEEE International Conference on Robotics and Automation (ICRA), May 29–June 3, Singapore. Reprinted with permission.

1. Introduction

High cost for time-consuming robot programming, performed by engineers, has become a key obstruction in industrial manufacturing. This has promoted the research toward faster and more intuitive means of robot programming, such as learning from demonstration, to which an introduction is presented in [Argall et al., 2009]. It is in this context desirable to make robot teaching available to a broader group of practitioners by minimizing the engineering work required during teaching of tasks.

A customary way to quickly mediate tasks to robots is to use lead-through programming, while saving trajectory data so that the robot can reproduce the motion. In this paper, the data are used to form dynamical movement primitives (DMPs). Early versions of these were presented in [Ijspeert et al., 2002], [Schaal et al., 2003] and [Ijspeert et al., 2003], and put into context in [Niekum et al., 2015]. Uncomplicated modification for varying tasks was emphasized in this literature. For example, the time scale was governed by one parameter, which could be adjusted to fit the purpose. Further, the desired final state could be adjusted, to represent a motion similar to the original one but to a different goal. DMPs applied on object handover with moving targets were addressed in [Prada et al., 2014]. The scalability in space was demonstrated in, *e.g.*, [Ijspeert et al., 2013].

The scenario considered in this paper is the unfavorable event that the last part of the motion generated by a certain DMP is unsatisfactory. There might be several reasons for this to occur. In the case where the starting points differ, the generated trajectory would still converge to the demonstrated end point, but take a modified path, where the modification would be larger for larger differences between the starting points. Further, the DMP might have been created in a slightly different setup, *e.g.*, for a different robot or robot cell. There might also have been a mistake in the teaching that the operator would have to undo. If the complete last part of the trajectory is of interest, it is not enough to modify the goal state only. One way to solve the problem would be to record an entirely new trajectory, and then construct a corresponding DMP. However, this would be unnecessarily time consuming for the operator, as only the last part of the trajectory has to be modified. Instead, the method described here allows the operator to lead the manipulator backwards, approximately along the part of the trajectory that should be adjusted, followed by a desired trajectory, as visualized in Fig. 1.

Hitherto, DMPs have usually been formed by demonstrations to get close to the desired behavior, followed by trajectory-based reinforcement learning, as presented in, *e.g.*, [Pastor et al., 2013; Kober et al., 2008; Abu-Dakka et al., 2015; Kroemer et al., 2010]. Compared to such refinements, the modification presented here is less time consuming and does not require engineering work. On the other hand, the previous work on reinforcement learning offers modulation based on sensor data, and finer movement adjustment. Therefore, the framework presented in this paper forms an intermediate step, where, if necessary, a DMP is modified to prepare for reinforcement learning, see Fig. 3. This modification can be used within a wide

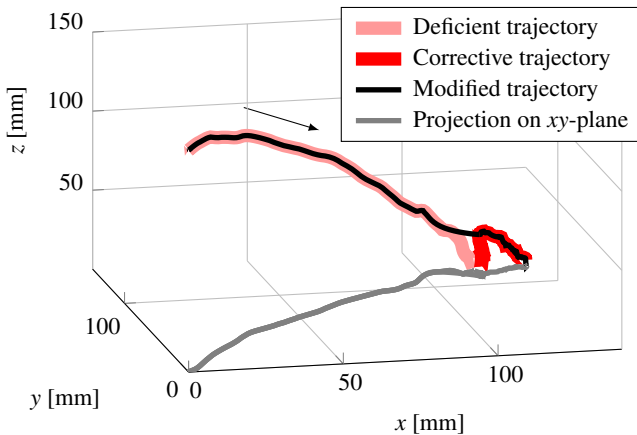


Figure 1. Trajectories of the robot’s end-effector from one of the experiments. The arrow indicates the motion direction. The deficient trajectory was generated from the original DMP. After that, the operator demonstrated the corrective trajectory. Merging of these, resulted in the modified trajectory. The projection on the xy -plane is only to facilitate the visualization.

range of tasks. In this paper, we exemplify by focusing on peg-in-hole tasks.

In [Pastor et al., 2013], online modulation, such as obstacle avoidance, was implemented for DMPs. This approach has been verified for several realistic scenarios, but requires an infrastructure for obstacle detection, as well as some coupling term parameters to be defined. It preserves convergence to the goal point, but since the path to get there is modified by the obstacle avoidance, it is not guaranteed to follow any specific trajectory to the goal. This is significant for, *e.g.*, a peg-in-hole task.

The paper is outlined as follows. Two example scenarios in which the framework would be useful are presented in Sec. 3, followed by a description of the method in Sec. 4. Experimental setup and results are described in Secs. 5 and 6, and finally a discussion and concluding remarks are presented in Secs. 7 and 8, respectively.

2. Problem Formulation

In this paper, we address the question whether it is possible to automatically interpret a correction, made by an operator, of the last part of a DMP trajectory, while still taking advantage of the first part. The human–robot interaction must be intuitive, and the result of a correction predictable enough for its purpose. The correction should result in a new DMP, of which the first part behaves qualitatively as the first part of the original DMP, whereas the last part resembles the last part of the corrective trajectory. Any discontinuity between the original and corrective trajectories must be mitigated.

3. Motivating Examples

We here describe two scenarios where the framework proves useful. These are evaluated in Secs. 5 and 6, where more details also are given.

3.1 Inadequate precision – Scenario A

Consider the setup shown in Fig. 2, where the button should be placed into the yellow case. A DMP was run for this purpose, but, due to any of the reasons described above, the movement was not precise enough, and the robot got stuck on its way to the target. Hitherto, such a severe shortcoming would have motivated the operator to teach a completely new DMP, and erase the old one. With the method proposed in this paper, the operator had the opportunity to approve the first part of the trajectory, and only had to modify the last part. This was done by leading the robot arm backwards, approximately along the faulty path, until it reached the acceptable part. Then, the operator continued to lead the arm along the desired path to the goal. When this was done, the acceptable part of the first trajectory was merged with the last part of the corrective trajectory. After that, a DMP was fitted to the resulting trajectory. Compared to just updating the target point, this approach also allowed the operator to determine the trajectory leading there. This scenario is referred to as Scenario A.

3.2 New obstacle – Scenario B

For the setup in Fig. 4, there existed a DMP for moving the robot arm from the right, above the button that was already inserted, to a position just above the hole in the leftmost yellow case. However, under the evaluation the operator realized that there would have been a collision if a button were already placed in the case in the middle. A likely reason for this to happen would be that the DMP was created in a slightly different scene, where the potential obstacle was not taken into account. Further, the operator desired to extend the movement to complete the peg-in-hole task, rather than stopping above the hole. With the method described herein, the action of the operator would be similar to that described in Sec. 3.1, again saving work compared to previous methods. This scenario is referred to as Scenario B.

4. Description of the Framework

In this section, the concept of DMPs is introduced. A method to determine what parts of the deficient and corrective trajectories to retain is presented, followed by a description of how these should be merged to avoid discontinuities. Finally, some implementation aspects are addressed. Figure 3 displays a schematic overview of the work flow of the application, from the user's perspective.

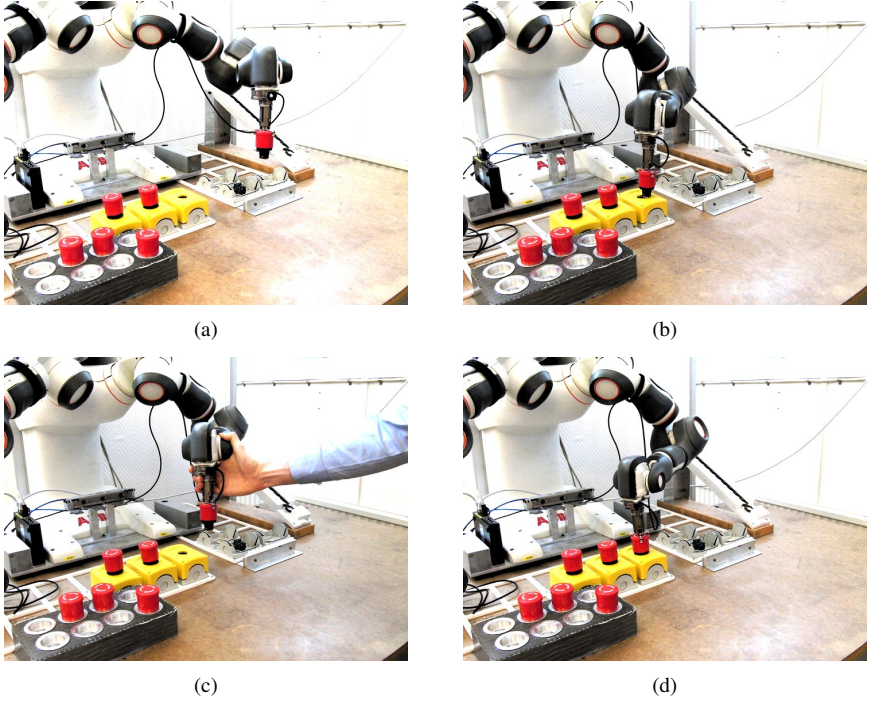


Figure 2. Scenario A. The evaluation started in (a), and in (b) the robot failed to place the button in the hole due to inadequate accuracy. Between (a) and (b), the deficient trajectory was recorded. The operator led the robot arm backwards (c), approximately along a proportion of the deficient trajectory, and subsequently led it to place the button properly, while the corrective trajectory was recorded. The robot then made the entire motion, starting in a configuration similar to that in (a), and ending as displayed in (d).

4.1 Dynamical movement primitives

A review of the DMP concept was presented in [Ijspeert et al., 2013], and here follows a short description of how it was applied in this paper. A certain trajectory, y , was modeled by the system

$$\tau \dot{y} = z \quad (4.1)$$

where z is determined by

$$\tau \dot{z} = \alpha_z (\beta_z (g - y) - z) + f(x) \quad (4.2)$$

In turn, $f(x)$ is a function given by

$$f(x) = \frac{\sum_{i=1}^{N_b} w_i \Psi_i(x)}{\sum_{i=1}^{N_b} \Psi_i(x)} x \cdot (g - y_0) \quad (4.3)$$

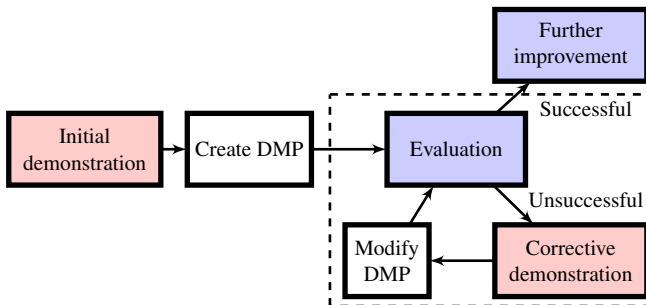


Figure 3. Schematic visualization of the work flow, from an operator’s perspective. A DMP was created based on a demonstration. Subsequently, the DMP was executed while evaluated by the operator. If unsuccessful, the operator demonstrated a correction, which yielded a modified DMP to be evaluated. Once successful, further improvement could be done by, *e.g.*, trajectory-based reinforcement learning, though that was outside the scope of this work. Steps that required direct, continuous interaction by the operator are marked with light red color. Steps that required some attention, such as supervision and initialization, are marked with light blue. The operations in the white boxes were done by the software in negligible computation time, and required no human involvement. The work in this paper focused on the steps within the dashed rectangle.

where the basis functions, $\Psi_i(x)$, take the form

$$\Psi_i(x) = \exp\left(-\frac{1}{2\sigma_i^2}(x - c_i)^2\right) \quad (4.4)$$

$$\tau\dot{x} = -\alpha_x x \quad (4.5)$$

Here, τ is a time constant, while α_z , β_z , and α_x are positive constants. Further, N_b is the number of basis functions, w_i is the weight for basis function i , y_0 is the starting point of the trajectory y , and g is the goal state; σ_i and c_i are the width and center of each basis function, respectively. Given a DMP, a robot trajectory can be generated from (4.1) and (4.2). Vice versa, given a demonstrated trajectory, y_{demo} , a corresponding DMP can be formed; g is then given by the end position of y_{demo} , whereas τ can be set to get a desired time scale. Further, the solution of a weighted linear regression problem in the sampled domain yields the weights

$$w_i = \frac{s^T \Gamma_i f_{\text{target}}}{s^T \Gamma_i s} \quad (4.6)$$

where

$$s = \begin{pmatrix} x^1(g - y_{\text{demo}}^1) \\ x^2(g - y_{\text{demo}}^1) \\ \vdots \\ x^N(g - y_{\text{demo}}^1) \end{pmatrix}, \quad \Gamma_i = \text{diag}(\Psi_i^1, \Psi_i^2 \dots \Psi_i^N) \quad (4.7)$$

$$f_{\text{target}} = \tau^2 \ddot{y}_{\text{demo}} - a_z(b_z(g - y_{\text{demo}}) - \tau \dot{y}_{\text{demo}}) = \begin{pmatrix} f_{\text{target}}^1 \\ f_{\text{target}}^2 \\ \vdots \\ f_{\text{target}}^N \end{pmatrix} \quad (4.8)$$

Here, N is the number of samples in the demonstrated trajectory.

4.2 Interpretation of corrective demonstration

If the evaluation of a trajectory was unsuccessful, a corrective demonstration and DMP modification should follow, as in Fig. 3. Denote by y_d the deficient trajectory, and by y_c the corrective one, of which examples are shown in Figs. 1, 5, and 6. A trajectory formed by simply appending y_c to y_d was likely to take an unnecessary detour. Thus, only the first part of y_d and the last part of y_c were retained. This is illustrated in Fig. 6. Denote by y_{cr} the retained part of the corrective trajectory. The operator signaled where to separate the corrective trajectory, during the corrective demonstration. In the current implementation, this was done by pressing a button in a terminal user interface, when the robot configuration corresponded to the desired starting point of y_{cr} , denoted y_{cr}^1 .

The next step was to determine which part of y_d to retain. This was chosen as the part previous to the sample of y_d that was closest to y_{cr}^1 , *i.e.*,

$$y_{dr}^m = y_d^m, \quad \forall m \in [1; M] \quad (4.9)$$

where

$$M = \underset{k=1 \dots K}{\text{argmin}} \, d(y_d^k, y_{cr}^1) \quad (4.10)$$

Here, d denotes distance, and K is the number of samples in y_d , see Fig. 5 for an illustration. The approach of using the shortest distance as a criterion, was motivated by the assumption that the operator led the robot arm back, approximately along the deficient trajectory, until the part that was satisfactory. At this point, the operator separated the corrective demonstration, thus defining y_{cr}^1 (see right marker in Fig. 6). By removing parts of the demonstrated trajectories, a significant discontinuity between the remaining parts was introduced. In order to counteract this, y_{dr} was modified into y_m , of which the following features were desired:

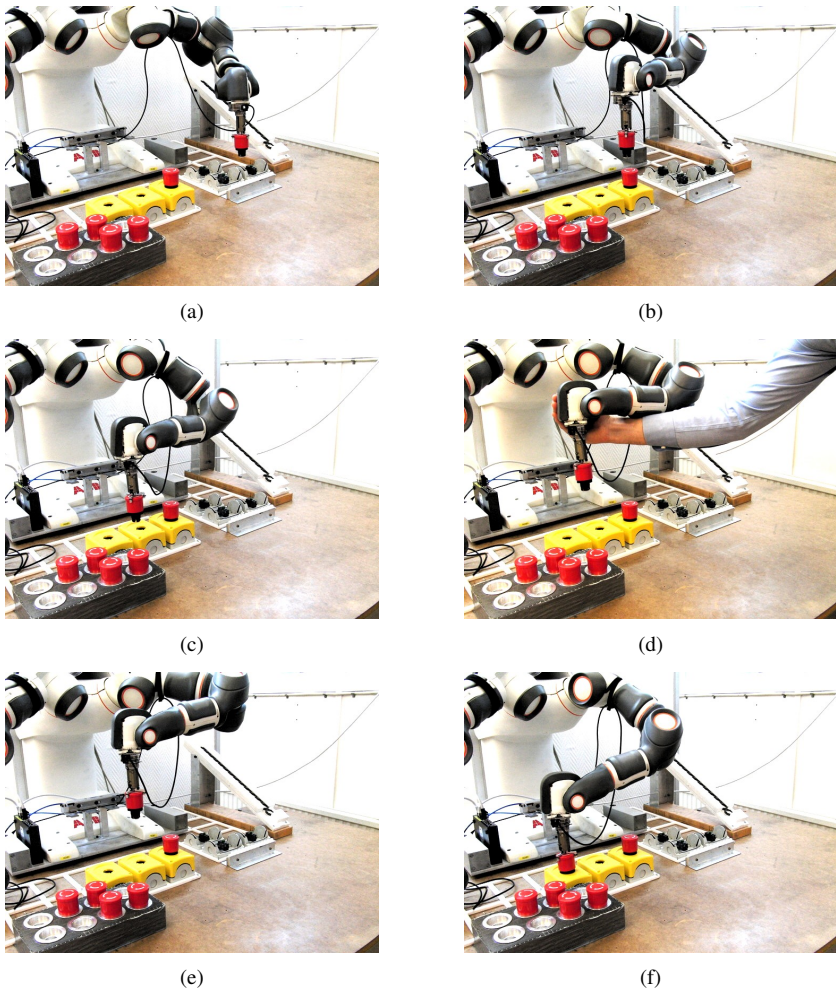


Figure 4. Scenario B. The initial goal was to move the button to the leftmost yellow case, above the hole, to prepare for placement. The evaluation started in (a), and in (b) the trajectory was satisfactory as the placed button was avoided. In (c), however, there would have been a collision if there was a button placed in the middle case. Further, it was desired to complete the peg-in-hole task, rather than stopping above the hole. Hence, the evaluated trajectory was considered deficient. In (d), the operator led the robot arm back, and then in a motion above the potential obstacle, and into the hole, forming the corrective trajectory. Based on the modified DMP, the robot started in a position similar to that in (a), avoided the potential obstacle in (e) and reached the new target in (f).

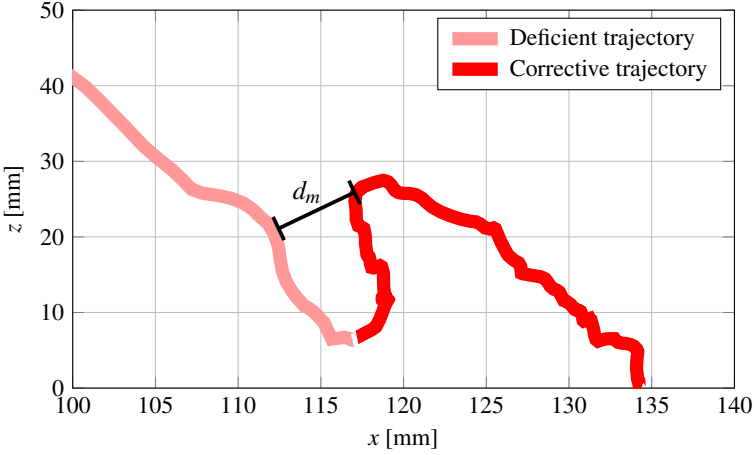


Figure 5. Visualization of shortest distance, here denoted d_m , used to determine the left separation marker in Fig. 6. The trajectories are the same as in Figs. 1 and 6, except that the modified trajectory is omitted.

- y_m should follow y_{dr} approximately;
- The curvature of y_m should be moderate;
- y_m should end where y_{cr} began, with the same movement direction in this point.

To find a suitable trade-off between these objectives, the following convex optimization problem was formulated and subsequently solved:

$$\underset{y_m}{\text{minimize}} \quad \|y_{dr} - y_m\|_2 + \lambda \|T_{(\Delta^2)} y_m\|_2 \quad (4.11)$$

$$\text{subject to} \quad y_m^M = y_{cr}^1 \quad (4.12)$$

$$y_m^M - y_m^{M-1} = y_{cr}^2 - y_{cr}^1 \quad (4.13)$$

Here, λ denotes a constant scalar, and $T_{(\Delta^2)}$ is a second-order finite difference operator. Thereafter, y_{cr} was appended on y_m , and one corresponding DMP was created, with the method described in the previous subsection. The next step in the workflow was to evaluate the resulting DMP, as shown in Fig. 3.

4.3 Software implementation

The research interface ExtCtrl [Blomdell et al., 2005; Blomdell et al., 2010], was used to send references to the low-level robot joint controller in the ABB IRC5 system [ABB Robotics, 2017a], at 250 Hz. Most of the programming was done in

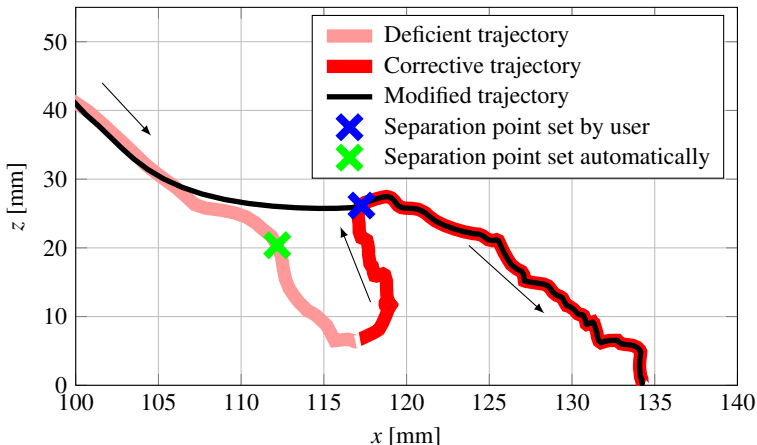


Figure 6. Same trajectories as in Fig. 1, but zoomed in on the corrective trajectory. Arrows indicate directions. The parts of the trajectories between the separation markers were not retained. The right, blue, separation point was determined explicitly by the operator during the corrective demonstration. The left, green, separation point was determined according to (4.10). Further, what was left of the deficient trajectory was modified for a smooth transition. However, the part of the corrective trajectory retained was not modified, since it was desired to closely follow this part of the demonstration. Note that the trajectories retained were not intended for direct play-back execution. Instead, they were used to form a modified DMP, which in turn generated a resulting trajectory, as shown in Figs. 8, 9 and 10.

C++, where DMPs were stored as objects. Among the data members of this class were the parameters τ , g and $w_{1\dots N_b}$, as well as some description of the context of the DMP and when it was created. It contained member functions for displaying the parameters, and for modifying g and τ . The communication between the C++ program and ExtCtrl was handled by the LabComm protocol [LabComm, 2017]. The C++ linear algebra library Armadillo [Sanderson and Curtin, 2016] was used in a major part of the implementation. Further, the code generator CVXGEN [Matingley and Boyd, 2012] was used to generate C code for solving the optimization problem in (4.11) to (4.13). By default, the solver code was optimized with respect to computation time. This resulted in a real-time application, in which the computation times were negligible in teaching scenarios. The optimization problem was typically solved well below one millisecond on an ordinary PC.

5. Experiments

The robot used in the experimental setup was a prototype of the dual-arm ABB YuMi [ABB Robotics, 2017b] (previously under the name FRIDA) robot, with 7

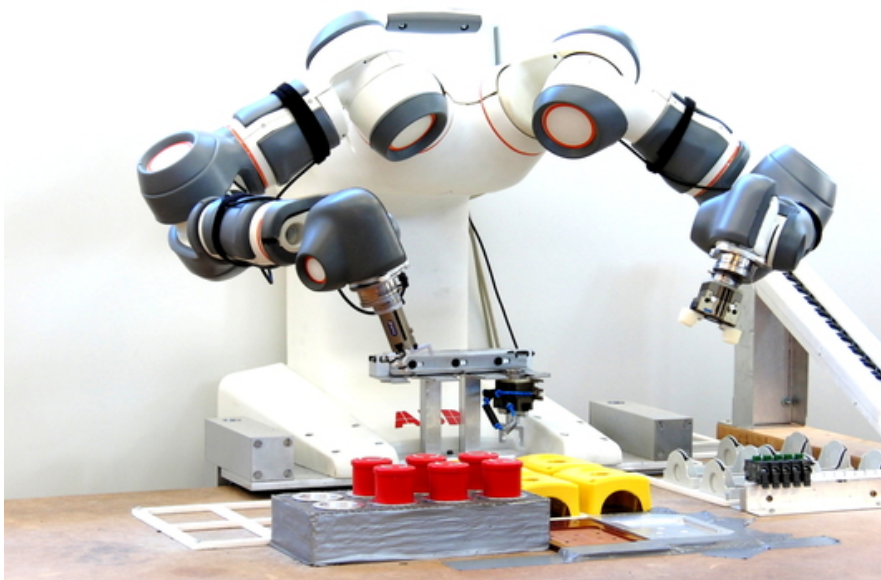


Figure 7. The ABB YuMi [ABB Robotics, 2017b] prototype robot used in the experiments.

joints per arm, see Fig. 7. The experiments were performed in real-time using the implementation described in Sec. 4.3. The computations took place in joint space, and the robot’s forward kinematics were used for visualization in Cartesian space in the figures presented. The scenarios in Sec. 3 were used to evaluate the proposed method. For each trial, the following steps were taken:

- An initial trajectory was taught, deliberately failing to meet the requirements, as explained in Sec. 3;
- Based on this, a DMP was created;
- The DMP was used to generate a trajectory similar to the initial one. This formed the deficient trajectory;
- A corrective trajectory was recorded;
- Based on the correction, a resulting DMP was formed automatically;
- The resulting DMP was executed for experimental evaluation.

First, Scenario A was set up for evaluation, see Sec. 3.1 and Fig. 2. The scenario started with execution of a deficient trajectory. For each attempt, a new deficient trajectory was created and modified. A total of 50 attempts were made.

Similarly, Scenario B (see Sec. 3.2 and Fig. 4) was set up, and again, a total of 50 attempts were made.

A video is available as a publication attachment, to facilitate understanding of the experimental setup and results. A version with higher resolution is available on [Karlsson, 2017].

6. Results

For each attempt of Scenario A, the robot was able to place the button properly in the yellow case after the modification. Results from two of these attempts are shown in Figs. 8 and 9. In the first case, the deficient trajectory went past the goal, whereas in the second case, it did not reach far enough.

Each of the attempts of Scenario B was also successful. After modification, the DMPs generated trajectories that moved the grasped stop button above the height of potential obstacles, in this case other stop buttons, and subsequently inserted it into the case. The result from one attempt is shown in Fig. 10.

7. Discussion

The subsequent step in this work is to integrate the presented framework with trajectory-based reinforcement learning [Pastor et al., 2013; Stulp et al., 2012],

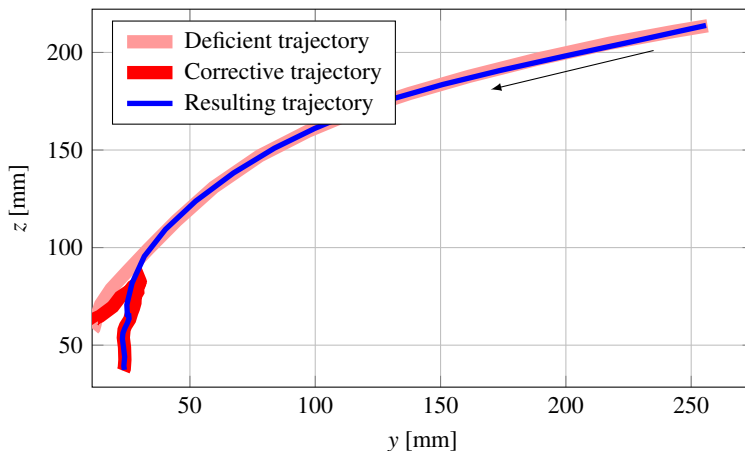


Figure 8. Trajectories from the experimental evaluation of Scenario A. The deficient trajectory went past the goal in the negative y -direction, preventing the robot from lowering the button into the hole. After correction, the robot was able to reach the target as the modified DMP generated the resulting trajectory.

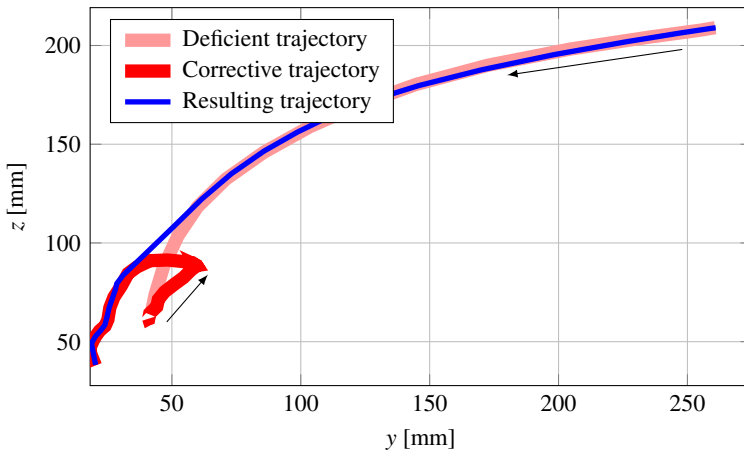


Figure 9. Similar to Fig. 8, except that in this case, the deficient trajectory did not reach far enough in the negative y -direction.

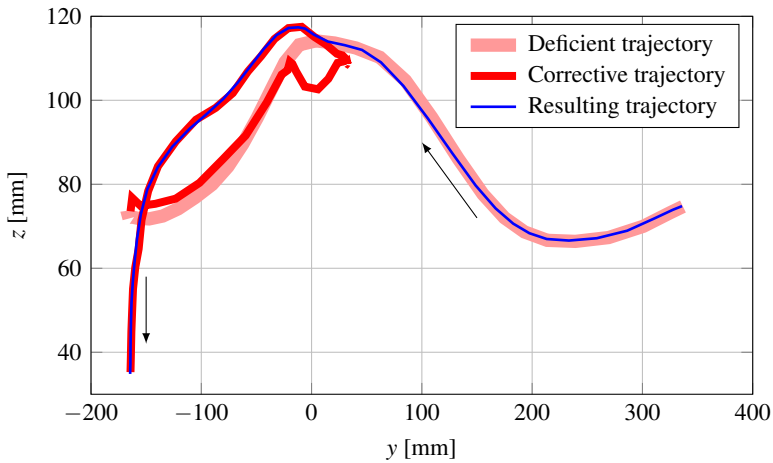


Figure 10. Trajectories from experimental evaluation of Scenario B. The deficient trajectory was lowered too early, causing a potential collision. After the correction, the robot was able to reach the target while avoiding the obstacles. The movement was also extended to perform the entire peg-in-hole task, rather than stopping above the hole.

in order to optimize the motion locally with respect to criteria such as execution time. The program should also be augmented to take the purpose of, and relation between, different DMPs into consideration. This extension will emphasize the necessity of keeping track of different states within the work flow. To this purpose, a state machine implemented in, *e.g.*, JGrafchart [Theorin, 2014], or the framework of behavior trees, applied on robot control in [Marzinotto et al., 2014], would be suitable. Extending the user interface with support for natural language, would possibly make this framework more user friendly.

Performing the computations in joint space instead of Cartesian space allowed the operator to determine the entire configuration of the 7 DOF robot arm, rather than the pose of the tool only. However, one could think of situations where the operator is not concerned by the configuration, and the pose of the tool would be more intuitive to consider. It would therefore be valuable if it could be determined whether the operator aimed to adjust the configuration or just the pose of the tool. For example, a large configuration change yielding a small movement of the tool, should promote the hypothesis that the operator aimed to adjust the configuration.

It should be stated that the scenarios evaluated here are not covering the whole range of plausible scenarios related to this method, and it remains as future work to investigate the generalizability, and user experience, more thoroughly. The last part of the resulting movement is guaranteed to follow the retained part of the corrective demonstration accurately, given enough DMP basis functions. Hence, the only source of error on that part is a faulty demonstration. For instance, the movement might require higher accuracy than what is possible to demonstrate using lead-through programming. Another limitation with this method is that it is difficult for the operator to very accurately determine which part of the faulty trajectory to retain, since this is done autonomously. However, for the experiments performed here, the estimation of the operator was sufficient to demonstrate the desired behavior. The benefit with this approach is that it saves time as the operator does not have to specify all details explicitly.

8. Conclusion

In this paper, an approach for modification of DMPs, using lead-through programming, was presented. It allowed a robot operator to modify the last part of a faulty generated trajectory, instead of demonstrating a new one from the beginning. Based on the corrective demonstration, modified DMPs were formed automatically. A real-time application, that did not require any additional engineering work by the user, was developed, and verified experimentally. A video showing the functionality is available as a publication attachment, and a version with higher resolution is available on [Karlsson, 2017].

Acknowledgments

The authors would like to thank Fredrik Bagge Carlson, Björn Olofsson and Karl Johan Åström at the Department of Automatic Control, Lund University, as well as Maj Stenmark, Mathias Haage and Jacek Malec at Computer Science, Lund University, for valuable discussions throughout this work. The authors are members of the LCCC Linnaeus Center and the ELLIIT Excellence Center at Lund University. The research leading to these results has received funding from the European Commission’s Framework Programme Horizon 2020 – under grant agreement No 644938 – SARAFun.

References

- ABB Robotics (2017a). *ABB IRC5*. URL: <http://new.abb.com/products/robotics/controllers/irc5> (visited on 2017-03-06).
- ABB Robotics (2017b). *ABB YuMi*. URL: <http://new.abb.com/products/robotics/yumi> (visited on 2016-01-23).
- Abu-Dakka, F. J., B. Nemeč, J. A. Jørgensen, T. R. Savarimuthu, N. Krüger, and A. Ude (2015). “Adaptation of manipulation skills in physical contact with the environment to reference force profiles”. *Autonomous Robots* **39**:2, pp. 199–217.
- Argall, B. D., S. Chernova, M. Veloso, and B. Browning (2009). “A survey of robot learning from demonstration”. *Robotics and Autonomous Systems* **57**:5, pp. 469–483.
- Blomdell, A., G. Bolmsjö, T. Brogårdh, P. Cederberg, M. Isaksson, R. Johansson, M. Haage, K. Nilsson, M. Olsson, T. Olsson, A. Robertsson, and J. Wang (2005). “Extending an industrial robot controller-implementation and applications of a fast open sensor interface”. *IEEE Robotics & Automation Magazine* **12**:3, pp. 85–94.
- Blomdell, A., I. Dressler, K. Nilsson, and A. Robertsson (2010). “Flexible application development and high-performance motion control based on external sensing and reconfiguration of abb industrial robot controllers”. In: *IEEE International Conference on Robotics and Automation (ICRA)*. May 3–8, Anchorage, Alaska, pp. 62–66.
- Ijspeert, A., J. Nakanishi, and S. Schaal (2003). “Learning control policies for movement imitation and movement recognition”. In: *Neural Information Processing System (NIPS)*. Vol. 15. December 5–10, Stateline, Nevada, pp. 1547–1554.
- Ijspeert, A. J., J. Nakanishi, H. Hoffmann, P. Pastor, and S. Schaal (2013). “Dynamical movement primitives: learning attractor models for motor behaviors”. *Neural Computation* **25**:2, pp. 328–373.

- Ijspeert, A. J., J. Nakanishi, and S. Schaal (2002). “Movement imitation with non-linear dynamical systems in humanoid robots”. In: *IEEE International Conference on Robotics and Automation (ICRA)*. Vol. 2. May 11–15, Washington, DC, pp. 1398–1403.
- Karlsson, M. (2017). *Modification of dynamical movement primitives*. Dept. Automatic Control, Lund University. URL: <https://www.youtube.com/watch?v=q998JUwofX4&feature=youtu.be> (visited on 2017-01-25).
- Kober, J., B. Mohler, and J. Peters (2008). “Learning perceptual coupling for motor primitives”. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. September 22–26, Nice, France, pp. 834–839.
- Kroemer, O., R. Detry, J. Piater, and P. J (2010). “Combining active learning and reactive control for robot grasping”. *Robotics and Autonomous Systems (RAS)* **58**:9, pp. 1105–1116.
- LabComm (2017). *Research tools and software*. Dept. Automatic Control, Lund University. URL: <http://www.control.lth.se/Research/tools.html> (visited on 2017-03-06).
- Marzinotto, A., M. Colledanchise, C. Smith, and P. Ogren (2014). “Towards a unified behavior trees framework for robot control”. In: *IEEE International Conference on Robotics and Automation (ICRA)*. May 31–June 7, Hong Kong, China, pp. 5420–5427.
- Mattingley, J. and S. Boyd (2012). “CVXGEN: a code generator for embedded convex optimization”. *Optimization and Engineering* **13**:1, pp. 1–27.
- Niekum, S., S. Osentoski, G. Konidaris, S. Chitta, B. Marthi, and A. G. Barto (2015). “Learning grounded finite-state representations from unstructured demonstrations”. *The International Journal of Robotics Research* **34**:2, pp. 131–157.
- Pastor, P., M. Kalakrishnan, F. Meier, F. Stulp, J. Buchli, E. Theodorou, and S. Schaal (2013). “From dynamic movement primitives to associative skill memories”. *Robotics and Autonomous Systems* **61**:4, pp. 351–361.
- Prada, M., A. Remazeilles, A. Koene, and S. Endo (2014). “Implementation and experimental validation of dynamic movement primitives for object handover”. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. September 14–18, Chicago, Illinois, pp. 2146–2153.
- Sanderson, C. and R. Curtin (2016). “Armadillo: a template-based C++ library for linear algebra”. *Journal of Open Source Software* **1**, p. 26.
- Schaal, S., A. Ijspeert, and A. Billard (2003). “Computational approaches to motor learning by imitation”. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **358**:1431, pp. 537–547.

- Stulp, F., J. Buchli, A. Ellmer, M. Mistry, E. A. Theodorou, and S. Schaal (2012). “Model-free reinforcement learning of impedance control in stochastic environments”. *IEEE Transactions on Autonomous Mental Development* **4**:4, pp. 330–341.
- Theorin, A. (2014). *A Sequential Control Language for Industrial Automation*. PhD thesis. TFRT–1104–SE, Dept. Automatic Control, Lund University, Lund, Sweden.

Paper II

Two-Degree-of-Freedom Control for Trajectory Tracking and Perturbation Recovery during Execution of Dynamical Movement Primitives

**Martin Karlsson Fredrik Bagge Carlson Anders Robertsson
Rolf Johansson**

Abstract

Modeling of robot motion as dynamical movement primitives (DMPs) has become an important framework within robot learning and control. The ability of DMPs to adapt online with respect to the surroundings, *e.g.*, to moving targets, has been used and developed by several researchers. In this work, a method for handling perturbations during execution of DMPs on robots was developed. Two-degree-of-freedom control was introduced in the DMP context, for reference trajectory tracking and perturbation recovery. Benefits compared to the state of the art were demonstrated. The functionality of the method was verified in simulations and in real-world experiments.

Accepted for the IFAC 2017 World Congress, July 9–14, Toulouse, France.
Reprinted with permission.

1. Introduction

Industrial robots have mostly operated in structured, predictable, environments through sequential execution of predefined motion trajectories. This implies high cost for engineering work, consisting of robot programming and careful work-space preparation. It also limits the range of tasks that are suitable for robots. Improving their ability to operate in unstructured environments with unforeseen events is therefore an important field of research.

This has motivated the development of dynamical movement primitives (DMPs), that are used to model and execute trajectories with an emphasis on online modification. Early forms were presented in [Ijspeert et al., 2002; Ijspeert et al., 2003; Schaal et al., 2000], and a review can be found in [Ijspeert et al., 2013]. The framework has been widely used by robot researchers. For instance, the ability to generalize demonstrated trajectories toward new, although static, goal positions has been used in [Niekum et al., 2015]. Online modulation with respect to a moving goal has been applied in [Prada et al., 2014] for object handover. A method to modify DMP parameters by demonstration has been presented in [Karlsson et al., 2017]. Learning and adaptation based on force/torque measurements has been explored in, *e.g.*, [Abu-Dakka et al., 2015; Pastor et al., 2013]. Previous work on DMP perturbation recovery in particular is elaborated on in Sec. 2.2.

In the standard form, without temporal coupling, a DMP would continue its time evolution regardless of any significant perturbation, as discussed in [Ijspeert et al., 2013]. Therefore, its behavior after the perturbation would likely be undesirable and not intuitive.

The research described in this paper addressed perturbation recovery for DMPs, and a method was developed where a two-degree-of-freedom controller was integrated with the DMP framework, see, *e.g.*, [Åström and Wittenmark, 2013] for an introduction to the two-degree-of-freedom control structure. The feedforward part of the controller promoted tracking of the DMP trajectory in the absence of significant perturbations, thus mitigating unnecessarily slow trajectory evolution due to temporal coupling acting on small tracking errors. The feedback part suppressed significant errors. The functionality of this method was verified in simulations, as well as in experiments in a real-time robot application. The robot used for experimental evaluation is shown in Fig. 1.

A code example is available on [Karlsson, 2017], to allow exploration of the system proposed. The system was also integrated in the Julia DMP package on [Bagge Carlson, 2016], originally based on [Ijspeert et al., 2013].

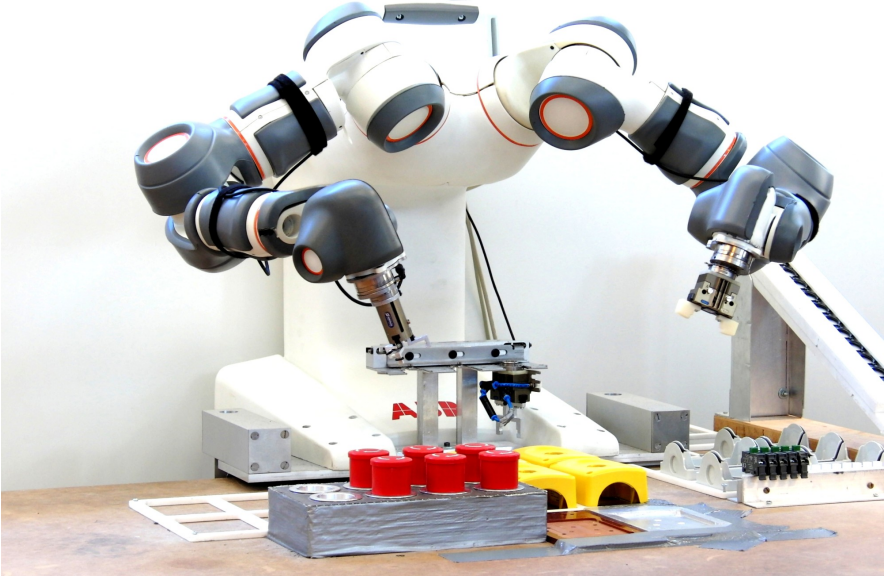


Figure 1. The ABB YuMi robot prototype used in the experiments, [ABB Robotics, 2017b].

2. Preliminaries

2.1 Dynamical movement primitives

A review of the DMP concept for robotics has been presented in [Ijspeert et al., 2013], and here follows a condensed description of the fundamentals. A trajectory, y , is modeled by the system

$$\tau^2 \ddot{y} = \alpha_z (\beta_z (g - y) - \tau \dot{y}) + f(x) \quad (4.1)$$

Here, τ is a time constant, α_z , β_z and α_x are positive constants, and x is a scalar phase parameter that evolves as

$$\tau \dot{x} = -\alpha_x x \quad (4.2)$$

Equation (4.1) is commonly written in the following equivalent form.

$$\tau \dot{z} = \alpha_z (\beta_z (g - y) - z) + f(x) \quad (4.3)$$

$$\tau \dot{y} = z \quad (4.4)$$

In (4.1) and (4.3), $f(x)$ is given by

$$f(x) = \frac{\sum_{i=1}^{N_b} w_i \Psi_i(x)}{\sum_{i=1}^{N_b} \Psi_i(x)} x \cdot (g - y_0) \quad (4.5)$$

where the basis functions, $\Psi_i(x)$, are determined as

$$\Psi_i(x) = \exp\left(-\frac{1}{2\sigma_i^2}(x-c_i)^2\right) \quad (4.6)$$

Here, N_b is the number of basis functions, w_i is the weight for basis function i , y_0 is the starting point of the trajectory y , and g denotes the goal state; σ_i and c_i are the width and center of each basis function, respectively. Based on the dynamical system in (4.3) and (4.4), a robot trajectory could be generated. Vice versa, given a demonstrated trajectory, y_{demo} , a corresponding DMP could be formed. The goal point g would then be given by the end position of y_{demo} , whereas τ could be set to get a desired time scale. Further, the weights could be determined by, *e.g.*, locally weighted linear regression, see [Atkeson et al., 1997; Schaal and Atkeson, 1998], with the solution

$$w_i = \frac{s^T \Gamma_i f_{\text{target}}}{s^T \Gamma_i s} \quad (4.7)$$

where

$$s = \begin{pmatrix} x^1(g - y_{\text{demo}}^1) \\ x^2(g - y_{\text{demo}}^1) \\ \vdots \\ x^N(g - y_{\text{demo}}^1) \end{pmatrix} \quad (4.8)$$

$$\Gamma_i = \text{diag}(\Psi_i^1, \Psi_i^2, \dots, \Psi_i^N) \quad (4.9)$$

$$f_{\text{target}} = \tau^2 \ddot{y}_{\text{demo}} - a_z(b_z(g - y_{\text{demo}}) - \tau \dot{y}_{\text{demo}}) \quad (4.10)$$

Here, N is the number of samples in the demonstrated trajectory.

2.2 Related work on DMP perturbation recovery

We here consider the case where a disturbance is introduced, such that the actual trajectory, denoted y_a , evolves differently from y , where y evolves according to (4.1) to (4.6). Without any coupling terms, the time evolution of (4.2) and (4.5) would be unaffected by a perturbation. This behavior is undesired, since it is then likely that the actual trajectory y_a deviates significantly from the intended trajectory even after the cause of the perturbation has vanished. This is more thoroughly described in [Ijspeert et al., 2002; Ijspeert et al., 2013]. To mitigate this problem, the solution described in the following paragraph has been suggested in [Ijspeert et al., 2013].

The following coupling terms were introduced.

$$\dot{e} = \alpha_e(y_a - y_c - e) \quad (4.11)$$

$$C_t = k_t e \quad (4.12)$$

$$\tau_a = 1 + k_c e^2 \quad (4.13)$$

Here, α_e , k_t and k_c are constant parameters. The parameter τ_a was used to determine the evolution rate of the entire dynamical system. Further, the term C_t was added to (4.3) so that the coupled version of y , denoted y_c , fulfilled the following.

$$\tau_a \dot{z} = \alpha_z (\beta_z (g - y_c) - z) + f(x) + C_t \quad (4.14)$$

$$\tau_a \dot{y}_c = z \quad (4.15)$$

A PD controller, given by

$$\ddot{y}_r = K_p (y_c - y_a) + K_v (\dot{y}_c - \dot{y}_a) \quad (4.16)$$

was used to drive y_a to y . Here, \ddot{y}_r denotes the reference acceleration, while K_p and K_v are control gains.

This approach from previous research has taken several important parts of disturbance recovery into account, and it should be emphasized that it forms the foundation of this presented work. In this section, however, some aspects are considered where there is room for improvement.

Denote by y_u an unperturbed trajectory generated by an uncoupled DMP, as described in Sec. 2.1. It is desirable that, in the absence of significant perturbations, y_a should follow y_c closely. If this would not be achieved, in addition to the deviation itself, y_a and y_c would be slowed down, compared to y_u , due to the temporal coupling in (4.13). This phenomenon is visualized in Fig. 5. In [Ijspeert et al., 2013], very high controller gains for (4.16) were suggested, which would have mitigated the issue under ideal conditions and unlimited magnitude of the control signals. Specifically, $K_p = 1000$ and $K_v = 125$ were chosen. However, even for moderate perturbations, this would imply control signals too large to be realized practically. For instance, a position error in Cartesian space of 1 dm would yield $\ddot{y}_r = 100 \text{ m/s}^2$. In Figs. 2 and 3, two example scenarios are displayed; one where the actual movement was stopped, and one where it was moved away from the nominal path. The method described in [Ijspeert et al., 2013] was used for recovery, with prohibitively large values of \ddot{y}_r as a consequence. Moreover, this control system is sensitive to noise and has a dangerously low delay margin of 12 ms.

Feedforward control has been used in the DMP context previously, but then only for low-level joint control, with motor torque commands as control signals, see [Pastor et al., 2009; Park et al., 2008]. This control structure was also applied in the internal controller used in the implementation in this present paper, see Sec. 6. This inner control design should not be confused with the feedforward control described in Sec. 4, which operated outside the internal robot controller, and was used to determine the reference acceleration for the robot.

3. Problem Formulation

In this paper, we address the question of whether perturbations of DMPs could be recovered from, while fulfilling the following requirements. Only moderate control

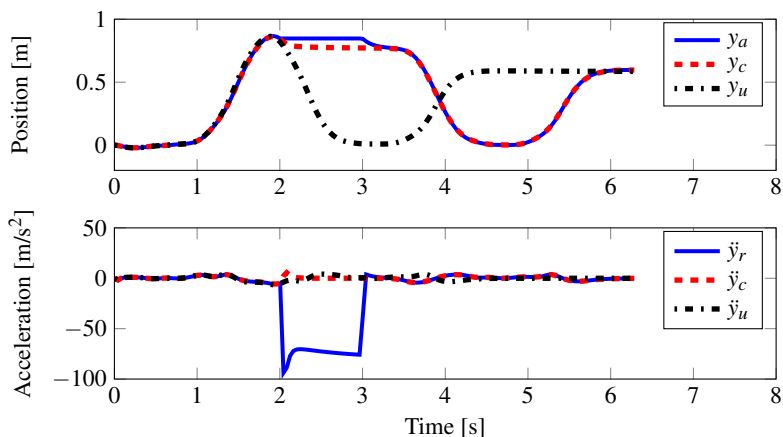


Figure 2. Simulated trajectories, where y_a was subjected to a stopping perturbation from 2 s to 3 s, using the approach in [Ijspeert et al., 2013]. When y_a was stopped, the evolution of y_c slowed down, and when y_a was released, it was driven to y_c and then behaved like a delayed version of y_u . This behavior was desired. However, a prohibitively large acceleration \ddot{y}_r was generated.

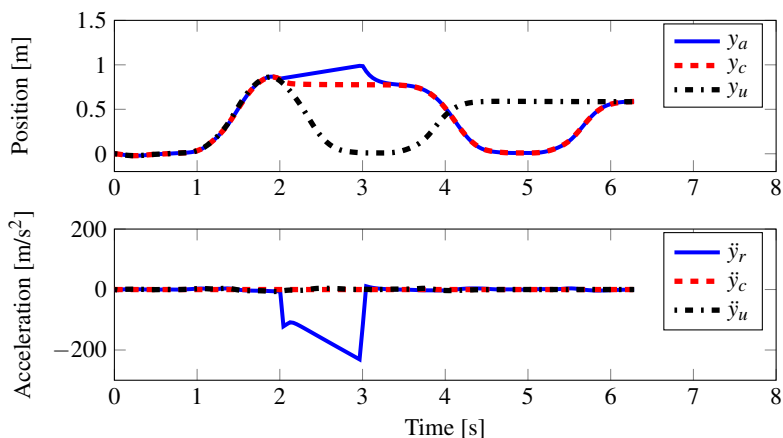


Figure 3. Similar to Fig. 2, except that y_a was moved away from the nominal path between 2 s to 3 s. Again, a prohibitively large acceleration \ddot{y}_r was generated.

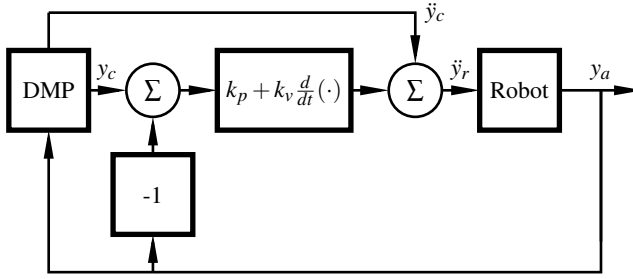


Figure 4. Schematic overview of the control structure described in Sec. 4. The block denoted 'Robot' includes the internal controller of the robot.

signals must be used. The benefits of the DMP framework described in [Ijspeert et al., 2013], *i.e.*, scalability in time and space as well as guaranteed convergence to the goal g , must be preserved. Further, in the absence of significant perturbations, the behavior of y_a should resemble that of the original DMP framework described in Sec. 2.1.

4. Method

Our proposed method extends that in [Ijspeert et al., 2013] as follows. The PD controller in (4.16) was augmented with feedforward control, as shown in (4.17). Further, the PD controller gains were moderate, to get a practically realizable control signal. Additionally, the time constant τ was introduced as a factor in the expression for the adaptive time parameter τ_a , see (4.13) and (4.18). Our method is detailed below.

In order for y_a to follow y_c , we applied the following control law.

$$\ddot{y}_r = k_p(y_c - y_a) + k_v(\dot{y}_c - \dot{y}_a) + \ddot{y}_c \quad (4.17)$$

Here, \ddot{y}_c was obtained by feedforwarding the acceleration of y_c . This allowed the controller to act also for zero position- and velocity errors. In turn, the trajectory tracking worked also for moderate controller gains; $k_p = 25$ and $k_v = 10$ are used throughout this paper. With these gains, the closed control loop had a double pole in -5 rad/s. Since the real parts were negative, the system was asymptotically stable, and since the imaginary parts were 0, it was critically damped. The delay margin was 130 ms, which was an improvement compared to 12 ms for the previous method, described in Sec. 2.2. A schematic overview of the control system is shown in Fig. 4.

Further, (4.13) was modified in order to include the nominal time constant τ , as follows.

$$\tau_a = \tau(1 + k_c e^2) \quad (4.18)$$

The coupling term C_t was omitted in this present method. This choice is elaborated on in Sec. 9.

Since τ_a was not constant over time, determining \ddot{y}_c was more involved than determining \dot{y} by differentiating (4.4). One option would be to approximate \ddot{y}_c by discrete-time differentiation of \dot{y}_c . However, instead we determined the instantaneous acceleration analytically as follows.

$$\ddot{y}_c = \frac{d}{dt}(\dot{y}_c) = \frac{d}{dt} \left(\frac{z}{\tau_a} \right) = \frac{\dot{z}\tau_a - z\dot{\tau}_a}{\tau_a^2} = \frac{\dot{z}\tau_a - 2\tau k_c z \dot{e}}{\tau_a^2} \quad (4.19)$$

where \dot{z} and \dot{e} are given by (4.11) and (4.14), respectively. It is noteworthy that the computation of \ddot{y}_c did not require any first- or second-order time-derivative of any measured signal, which would have required prior filtering to mitigate amplification of high-frequency noise. Similarly, \dot{y}_c was determined by (4.14) and (4.15). In contrast, the computation of \dot{y}_a was complemented with a low-pass filter, to mitigate amplification of measurement noise.

5. Simulations

Two different perturbations were considered in the following simulations; one where y_a was stopped, and one where it was moved. The perturbations took place from time 2 s to 3 s. The systems were sampled at 250 Hz. The same DMP, yielding the same y_u , was used in each trial. The adaptive time parameter τ_a was determined according to (4.18) in all simulations, to get comparable time scales. First, the controller detailed in [Ijspeert et al., 2013] was applied. Except for the perturbations themselves, the conditions were assumed to be ideal, *i.e.*, no delay and no noise were present. The results are shown in Figs. 2 and 3. Despite ideal conditions, prohibitively large accelerations were generated by the controller in both cases.

Figure 5 shows the result from a simulation where the controller detailed in [Ijspeert et al., 2013] was used, except that the gains were lowered to moderate values. The conditions were ideal, and no perturbation was present. This resulted in reasonable control signals. However, small control errors in combination with the temporal coupling slowed down the evolution of the coupled system as well as the actual movement.

Thereafter, the controller proposed in this paper, described in Sec. 4, was used. In order to verify robustness under realistic conditions, noise and time delay were introduced. Position measurement noise, and velocity process noise, were modeled as zero mean Gaussian white noise, with standard deviations of 1 mm and 1 mm/s, respectively. Further, an additional configuration dependent forward kinematics error was modeled as a slowly varying position measurement error with standard deviation 1 mm. The time delay between the process and the controller was $L = 12$ ms. This delay was suitable to simulate since it corresponds both to the delay margin of the method suggested in [Ijspeert et al., 2013], and to the actual delay in the im-

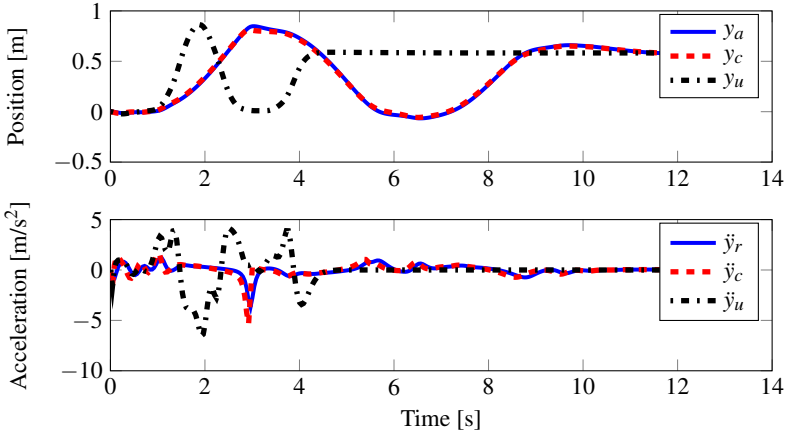


Figure 5. Simulation with the control structure in [Ijspeert et al., 2013], except that the gains were lower ($K_p = 25$ and $K_v = 10$). The reference acceleration was of reasonable magnitude, but the coupled and real systems were slowed down due to small tracking errors combined with temporal coupling.

plementation presented in this paper, see Sec. 6. (It is, however, a coincidence that these two have the same value. Nevertheless, this shows that a 12 ms delay margin is not necessarily enough.) The results are shown in Figs. 6 and 7. For comparison, the method in [Ijspeert et al., 2013], with the large gains, was also evaluated under these conditions, although without any perturbation except for the noise. Because of the time delay, this system was unstable, as shown in Fig. 8.

6. Implementation of Real-Time Application

The implementation presented here was performed on a prototype of the dual-arm ABB YuMi robot (previously under the name FRIDA), [ABB Robotics, 2017b], with 7 joints per arm, see Fig. 1. The method described in Sec. 4 was implemented in C++, and the linear algebra library Armadillo, see [Sanderson and Curtin, 2016], was used in a large proportion of the program. The research interface ExtCtrl, [Blomdell et al., 2005; Blomdell et al., 2010], was used to send references to the low-level robot joint controller in the ABB IRC5 system, [ABB Robotics, 2017a]. The LabComm protocol, [LabComm, 2017], was used to manage the communication between the C++ program and ExtCtrl. Similar to the simulations, the control system ran at 250 Hz, and the delay between process and controller was 3 sample periods, corresponding to 12 ms.

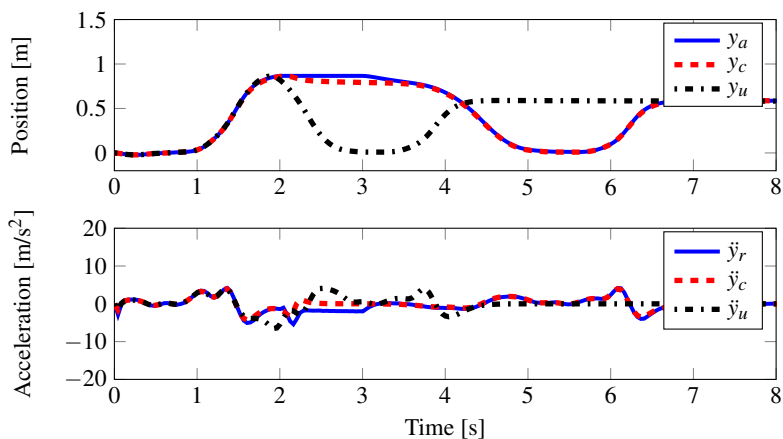


Figure 6. Similar to Fig. 2, but with modeled noise and delay, and using the controller presented in this paper. The behavior was satisfactory both regarding position and acceleration.

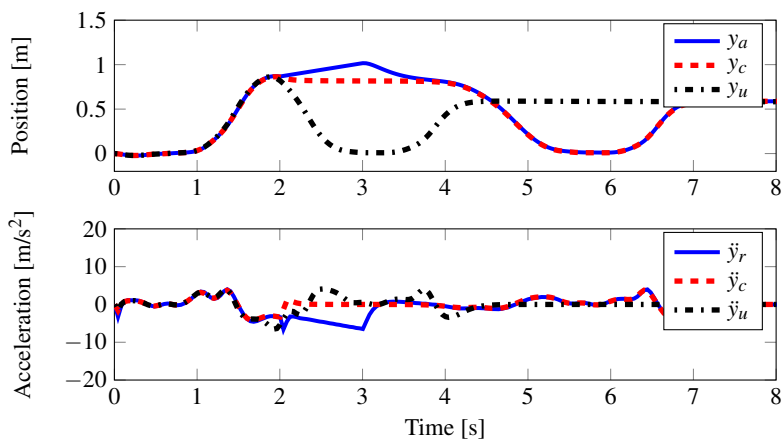


Figure 7. Similar to Fig. 6, except that y_a was moved away from the nominal path between 2 s to 3 s. Again, the behavior was satisfactory both regarding position and acceleration.

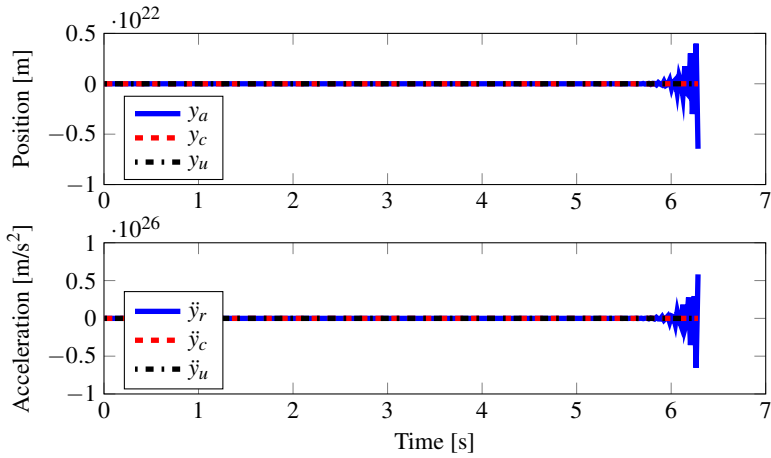


Figure 8. Using the control system in [Ijspeert et al., 2013], subject to the simulated noise and time delay, resulted in unstable behavior.

7. Experimental Setup

The real-time implementation described in Sec. 6 was used for evaluation. The computations took place in joint space, and the robot's forward kinematics were used for visualization in Cartesian space in the figures presented. The functionality of the method was evaluated in two assembly scenarios. The assembly parts used are shown in Fig. 9.

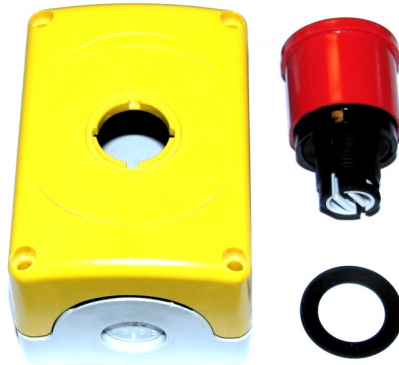


Figure 9. Yellow case (left), stop button (upper right) and gasket (lower right) used in the experiments.

For both scenarios, a new DMP for placing a stop button into the hole of a corresponding case had been taught to the robot by lead-through programming, based on [Stolt et al., 2015], prior to each trial. This implied some variation among the demonstrated trajectories, even though they were qualitatively similar. Subsequently, the DMP was executed on the robot. During the execution, a human perturbed the movement of the robot by physical contact. A wrist-mounted ATI Mini force/torque sensor was used to measure the contact force, and a proportional acceleration, in the same direction as the force, was added to \ddot{y}_r as a load disturbance.

In the first scenario, the human introduced two perturbations during the DMP execution. The first perturbation was formed by moving the end-effector away from its path, and then releasing it. The second perturbation consisted of a longer, unstructured, movement later along the trajectory.

In the second scenario, a human co-worker realized that the stop buttons in the current batch were missing rubber gaskets, and acted to modify the robot trajectory, allowing the co-worker to attach the gasket on the stop button manually. During execution of the DMP, the end-effector was stopped and lifted to a comfortable height by the co-worker. Thereafter, the gasket was attached, and finally the end-effector was released. For the sake of completeness, the modified trajectory was used to form yet another DMP, which allowed the co-worker to attach the gaskets without perturbing the trajectory of the robot, for the remaining buttons in the batch. To verify this functionality, one such modified DMP was executed at the end of each trial.

The first and second scenarios are visualized in Figs. 10 and 11, respectively. To verify repeatability, 50 similar trials were performed for each scenario.

8. Experimental Results

Data from a trial of the first scenario are displayed in Fig. 12. The two disturbances were successfully recovered from as intended. The reference acceleration was of reasonable magnitude. The results from all 50 trials were qualitatively mutually similar.

Data from a trial of the second scenario are displayed in Fig. 13. First, the perturbation was successfully recovered from as intended. The reference acceleration was of reasonable magnitude. When the modified DMP was executed, it behaved like a smooth version of the perturbed original trajectory. Again, the results from all 50 trials were qualitatively mutually similar.

To facilitate understanding of the experimental setup and results, a video is publicly available on [Karlsson, 2016].

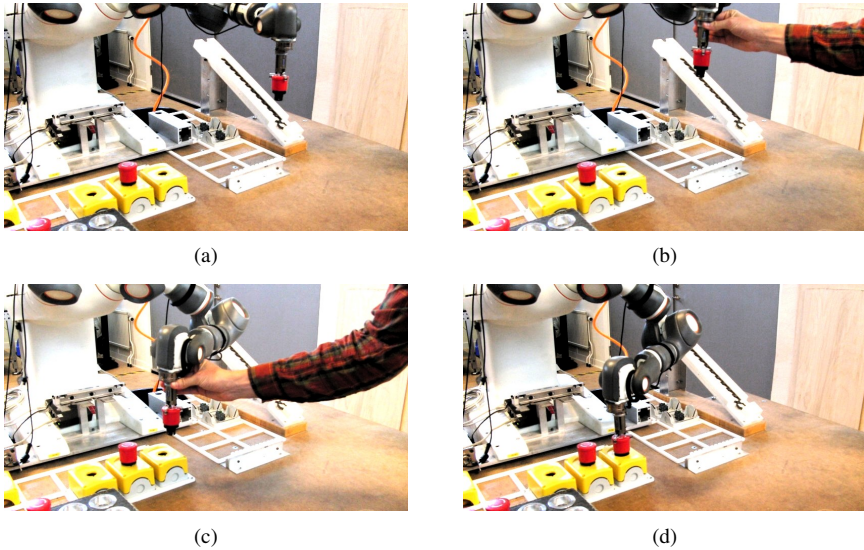


Figure 10. First scenario. In (a), the robot started to execute a DMP for placing the stop button in the rightmost yellow case. A human perturbed the motion twice. The first perturbation (b) was formed by moving the end-effector away from its path, and then releasing it. The second perturbation (c) lasted for a longer time, and consisted of unstructured movement. The robot recovered from both perturbations, and managed to place the stop button in the case (d). Data from one trial are shown in Fig. 12.

9. Discussion

Compared to previous related research, described in [Ijspeert et al., 2013], the method in this paper contained the following extensions. Feedforward control was added to the PD controller, thus forming a two-degree-of-freedom controller. Further, the PD controller gains were reduced to moderate magnitudes. The expression for τ_a was also modified, to include the nominal time constant τ as a factor. These changes resulted in the following benefits, compared to the previous method. The feedforward part allowed the controller to act also for insignificant position- and velocity error, thus improving the trajectory tracking. Because of this, the large controller gains used in [Ijspeert et al., 2013], that were used to mitigate significant tracking errors, could be reduced to moderate magnitudes. In turn, using moderate gains instead of very large ones, resulted in control signals that were practically realizable, instead of prohibitively large. It also improved the delay margin significantly. The aspects above form the main contribution of this paper. In contrast, the modification of the expression for τ_a was not the main focus of this paper, but it was necessary since it allowed the actual trajectory to converge to the trajectory defined

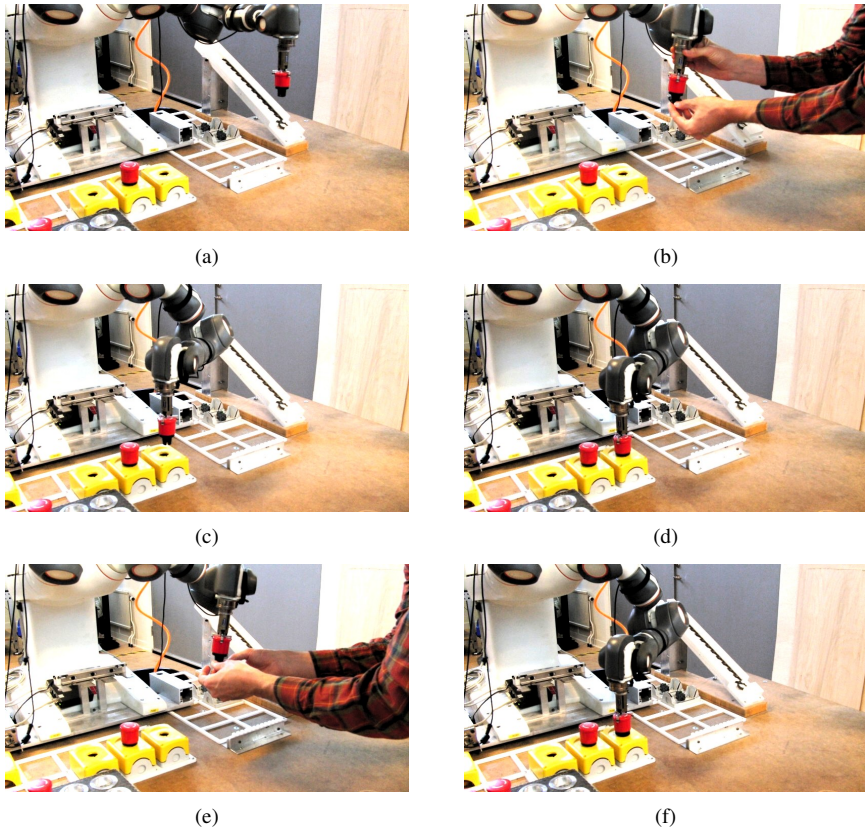


Figure 11. Second scenario. The robot started its motion toward the rightmost yellow case in (a). The end-effector was stopped and lifted, and the gasket was mounted in (b). The robot was then released, and continued its motion to the case, (c) and (d). The actual trajectory was saved and used to form a modified DMP, and the robot was reset to a configuration similar to that in (a). When executing the modified DMP, the human co-worker could attach the gasket without perturbing the motion of the robot (e). The robot finished the modified DMP in (f). Data from one trial are shown in Fig. 13.

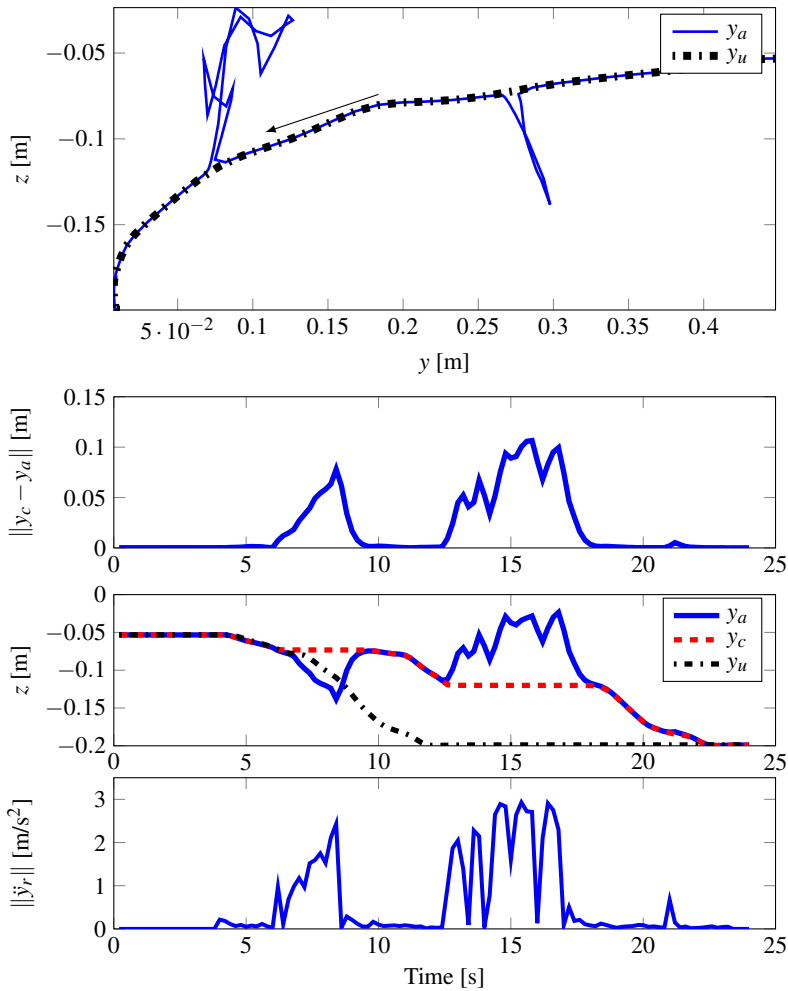


Figure 12. Experimental data from a trial of the first scenario. The first (from above) plot shows the path of the end-effector in the Cartesian base frame of the robot, projected on the yz -plane. The arrow indicates the movement direction, which started in the upper right and finished in the lower left of the plot. The two perturbations are clearly visible. The second plot shows the distance between y_a and y_c over time. In the third plot, it can be seen that the evolution of y_c slowed down during each perturbation. Subsequently, y_a recovered, and when it was close to y_c , the movement continued as a delayed version of y_u . The reference acceleration was of reasonable magnitude, as shown in the fourth plot.

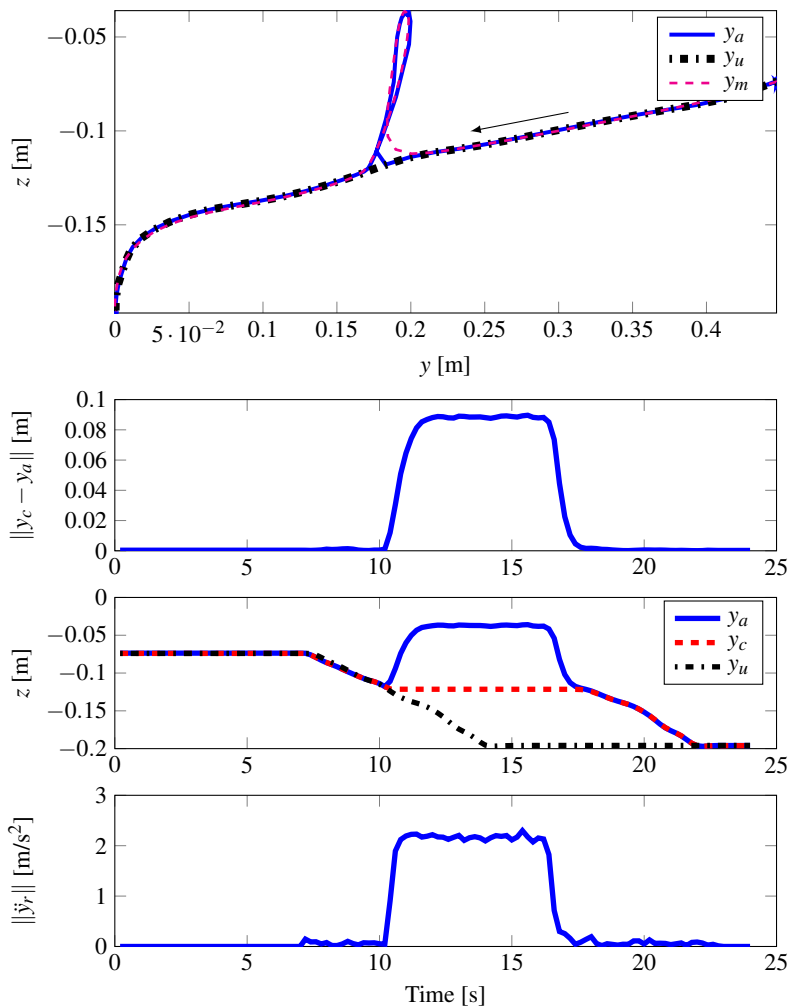


Figure 13. Experimental data from a trial of the second scenario. The organization of this figure is similar to that of Fig. 12. The perturbation for stopping and lifting the end-effector took place from time 10s to 17.5s, and is clearly visible in each plot. This perturbation was recovered from as intended, and the reference acceleration was of reasonable magnitude. The uppermost plot also displays the measured trajectory obtained by executing the modified DMP, denoted y_m . It behaved like a smooth version of the perturbed original trajectory y_a .

by the DMP, with time constant τ . Without this modification, the time parameter τ would not have affected the trajectory generated by the DMP. Instead, τ_a would have converged to 1, regardless of τ , which would not have been desirable.

The work presented here focused on the control structure for trajectory tracking and perturbation recovery, rather than on the perturbations themselves. Even though the perturbations in the experiments considered here emerged from physical contact with a human, the control structure would work similarly for any type of perturbation. There are many other possible perturbations, *e.g.*, a pause of the movement until a certain condition is fulfilled, superpositioned motion control signals to explore the surroundings with a force/torque sensor, a detour to allow line-of-sight between a camera and a part of the work-space, or any other unforeseen deviation from the reference trajectory defined by a DMP.

It is necessary to implement saturation on the control signals, in order to prevent too large acceleration and velocity for large perturbations. Such boundaries were implemented, but never reached in the experiments in this work.

The coupling term C_t has been introduced in previous research to drive y_c toward y_a when these were different, see Sec. 2.2. However, whether this effect is desired, and to what extent, is context dependent. Further, the effect would be mitigated by the temporal coupling, that would slow down the evolution of y_c in (4.14) and (4.15). Which of these effects that would be dominant in different cases would be difficult to predict intuitively. For these two reasons, the coupling term was not included in the method proposed here, though it would be straight forward to implement. It was, however, included in the simulations where the previous method, described in Sec. 2.2, was evaluated. During the perturbations in Figs. 2 and 3, the effect of the temporal coupling was dominant, as y_c did not approach y_a significantly.

Apart from the perturbations induced by the human, the motion of the robot was affected by process- and measurement noise. After applying the forward kinematics to determine the position of the end-effector, the accuracy was typically ± 1 mm. Furthermore, some movement might require higher precision than what would be possible to demonstrate using lead-through programming. Then, *e.g.*, teleoperation could be used for demonstration instead.

In the current implementation, the actual trajectory returned to the reference trajectory, approximately where it started to deviate. This might not always be desired. For instance, it might sometimes be more practical to connect further along the reference trajectory, *e.g.*, after avoiding an obstacle. A lower value of k_c would result in such behavior, however, it must then be known what value of k_c that should be used. Further, one could think of scenarios where it would not be desirable to connect to the reference trajectory, *e.g.*, if a human would modify the last part of the trajectory to a new end point. Hence, future work includes development of a method to determine the desired behavior after a perturbation.

The method presented in this paper would be useful for executing the desired behavior, once it could be determined. Nevertheless, one can think of various scenarios where the recovery presented here would be desirable, such as those in Sec. 7.

10. Conclusion

In this work, it was shown how perturbations of DMPs could be recovered from, while preserving the characteristics of the original DMP framework in the absence of significant perturbations. Feedforward control was used to track the reference trajectory generated by a DMP. Feedback control with moderate gains was used to suppress deviations. This design is the first, to the best of our knowledge, that takes the following aspects into account. In the absence of significant disturbances, the position error must be small enough, so that the dynamical system would not slow down unnecessarily due to the temporal coupling. Very large controller gains would result in small errors under ideal conditions, but are not practically realizable. On the other hand, if the gains are moderate and only feedback control is used, too large errors occur.

Feedforward allowed the controller to act even without significant error, which in turn allowed for moderate controller gains. The suggested method was verified in simulations, and a real-time application was implemented and evaluated, with satisfactory results. A video of the experiments is available on [Karlsson, 2016].

Acknowledgments

The authors would like to thank Björn Olofsson and Fredrik Magnusson at the Department of Automatic Control, Lund University, as well as Maj Stenmark, Mathias Haage and Jacek Malec at Computer Science, Lund University, for valuable discussions throughout this work. Anthony Remazeilles at Tecnalia, Donostia, and Diogo Almeida at KTH, Stockholm, are gratefully acknowledged for pointing out some of the previous research. The authors are members of the LCCC Linnaeus Center and the ELLIIT Excellence Center at Lund University. The research leading to these results has received funding from the European Commission's Framework Programme Horizon 2020 – under grant agreement No 644938 – SARAFun.

References

- ABB Robotics (2017a). *ABB IRC5*. URL: <http://new.abb.com/products/robotics/controllers/irc5> (visited on 2017-03-06).
- ABB Robotics (2017b). *ABB YuMi*. URL: <http://new.abb.com/products/robotics/yumi> (visited on 2016-01-23).

- Abu-Dakka, F. J., B. Nemeč, J. A. Jørgensen, T. R. Savarimuthu, N. Krüger, and A. Ude (2015). “Adaptation of manipulation skills in physical contact with the environment to reference force profiles”. *Autonomous Robots* **39**:2, pp. 199–217.
- Åström, K. J. and B. Wittenmark (2013). *Computer-Controlled Systems: Theory and Design*. Courier Corporation, Mineola, NY.
- Atkeson, C. G., A. W. Moore, and S. Schaal (1997). “Locally weighted learning for control”. In: *Lazy Learning*. Springer, Dordrecht, Netherlands, pp. 75–113.
- Bagge Carlson, F. (2016). *DynamicMovementPrimitives.jl*. Department of Automatic Control, Lund University. URL: <https://github.com/baggepinnen/DynamicMovementPrimitives.jl> (visited on 2017-03-21).
- Blomdell, A., G. Bolmsjö, T. Brogårdh, P. Cederberg, M. Isaksson, R. Johansson, M. Haage, K. Nilsson, M. Olsson, T. Olsson, A. Robertsson, and J. Wang (2005). “Extending an industrial robot controller-implementation and applications of a fast open sensor interface”. *IEEE Robotics & Automation Magazine* **12**:3, pp. 85–94.
- Blomdell, A., I. Dressler, K. Nilsson, and A. Robertsson (2010). “Flexible application development and high-performance motion control based on external sensing and reconfiguration of abb industrial robot controllers”. In: *IEEE International Conference on Robotics and Automation (ICRA)*. May 3–8, Anchorage, Alaska, pp. 62–66.
- Ijspeert, A., J. Nakanishi, and S. Schaal (2003). “Learning control policies for movement imitation and movement recognition”. In: *Neural Information Processing System (NIPS)*. Vol. 15. December 5–10, Stateline, Nevada, pp. 1547–1554.
- Ijspeert, A. J., J. Nakanishi, H. Hoffmann, P. Pastor, and S. Schaal (2013). “Dynamical movement primitives: learning attractor models for motor behaviors”. *Neural Computation* **25**:2, pp. 328–373.
- Ijspeert, A. J., J. Nakanishi, and S. Schaal (2002). “Movement imitation with nonlinear dynamical systems in humanoid robots”. In: *IEEE International Conference on Robotics and Automation (ICRA)*. Vol. 2. May 11–15, Washington, DC, pp. 1398–1403.
- Karlsson, M. (2016). *Experimental evaluation of DMP perturbation recovery*. Youtube. URL: <https://www.youtube.com/watch?v=u8GwsSsLOTI> (visited on 2016-10-04).
- Karlsson, M. (2017). *DMP perturbation, simulation example*. Department of Automatic Control, Lund University. URL: https://gitlab.control.lth.se/cont-mkr/dmp_perturbation_sim_example (visited on 2017-03-21).

- Karlsson, M., A. Robertsson, and R. Johansson (2017). “Autonomous interpretation of demonstrations for modification of dynamical movement primitives”. In: *IEEE International Conference on Robotics and Automation (ICRA)*. May 29–June 3, Singapore.
- LabComm (2017). *Research tools and software*. Dept. Automatic Control, Lund University. URL: <http://www.control.lth.se/Research/tools.html> (visited on 2017-03-06).
- Niekum, S., S. Osentoski, G. Konidaris, S. Chitta, B. Marthi, and A. G. Barto (2015). “Learning grounded finite-state representations from unstructured demonstrations”. *The International Journal of Robotics Research* **34**:2, pp. 131–157.
- Park, D.-H., H. Hoffmann, P. Pastor, and S. Schaal (2008). “Movement reproduction and obstacle avoidance with dynamic movement primitives and potential fields”. In: *Humanoids 2008—8th IEEE-RAS International Conference on Humanoid Robots*. December 1–3, Daejeon, Korea, pp. 91–98.
- Pastor, P., H. Hoffmann, T. Asfour, and S. Schaal (2009). “Learning and generalization of motor skills by learning from demonstration”. In: *IEEE International Conference on Robotics and Automation (ICRA)*. May 12–17, Kobe, Japan, pp. 763–768.
- Pastor, P., M. Kalakrishnan, F. Meier, F. Stulp, J. Buchli, E. Theodorou, and S. Schaal (2013). “From dynamic movement primitives to associative skill memories”. *Robotics and Autonomous Systems* **61**:4, pp. 351–361.
- Prada, M., A. Remazeilles, A. Koene, and S. Endo (2014). “Implementation and experimental validation of dynamic movement primitives for object handover”. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. September 14–18, Chicago, Illinois, pp. 2146–2153.
- Sanderson, C. and R. Curtin (2016). “Armadillo: a template-based C++ library for linear algebra”. *Journal of Open Source Software* **1**, p. 26.
- Schaal, S. and C. G. Atkeson (1998). “Constructive incremental learning from only local information”. *Neural Computation* **10**:8, pp. 2047–2084.
- Schaal, S., S. Kotosaka, and D. Sternad (2000). “Nonlinear dynamical systems as movement primitives”. In: *IEEE International Conference on Humanoid Robotics*. September 7–8, Boston, MA, pp. 1–11.
- Stolt, A., F. B. Carlson, M. G. Ardakani, I. Lundberg, A. Robertsson, and R. Johansson (2015). “Sensorless friction-compensated passive lead-through programming for industrial robots”. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. September 28–October 2, Hamburg, Germany, pp. 3530–3537.

Paper III

Detection of Contact Force Transients during Robotic Assembly without a Force Sensor

Martin Karlsson Anders Robertsson Rolf Johansson

Abstract

In this research, robot joint torques are used to recognize contact force transients induced by snap-fit assembly, thus detecting when the task is completed. The approach does not assume any external sensor, which is a benefit compared to the state of the art. The joint torque data is used as input to a recurrent neural network (RNN), and the output of the RNN indicates whether a snap-fit has occurred or not. A real-time application for snap-fit detection is developed, and verified experimentally on an industrial robot.

1. Introduction

In the context of robotic assembly, robots commonly make series of movements, and switch between these when certain criteria are fulfilled. Such criteria usually consist of thresholds on measured signals, *e.g.*, positions and contact forces.

In this paper, a method to detect snap-fits during robotic assembly is presented and evaluated. In [Stolt et al., 2015], this was achieved by detecting contact force transients induced by the snap-fit, using a force/torque sensor. This detection reduced the assembly time, compared to using a force threshold. It also removed the necessity to determine any level of the force threshold, which would have required considerable engineering work and explicit programming of the robot.

Here, we continue the work presented in [Stolt et al., 2015], with the following extensions. In [Stolt et al., 2015], a force/torque sensor was used to measure the contact force/torque. Such sensors and systems are usually expensive, with costs comparable to the robot itself. If attached to the wrist of the robot, it would introduce extra weight that the robot would have to lift and move. Further, some robot models do not support any seamless attachment of such sensors. If the force sensor would be attached to an object in the work space, *e.g.*, a table, this would imply restrictions on where the assembly could take place. In this work, we use robot joint torque measurements for the detection, thus avoiding the requirement of a force/torque sensor. This introduces a new difficulty; due to friction in the robot joints, some information is lost when using joint torques compared to a force/torque sensor. In [Stolt et al., 2015], a support vector machine (SVM) was used for classification, whereas in the present approach, a recurrent neural network (RNN), which is an artificial neural network specialized in processing sequential data, was used. The RNN was implemented and trained using TensorFlow [Abadi et al., 2016; TensorFlow, 2017], a software library for numerical computation.

Machine learning for analyzing contact forces in robotic assembly was also applied in [Rodriguez et al., 2010], where force measurements were used as input to an SVM, to distinguish between successful and failed assemblies. A verification system, specialized in snap-fit assembly, was developed in [Rojas et al., 2012]. Similar to [Stolt et al., 2015] and [Rodriguez et al., 2010], a force/torque sensor was assumed in [Rojas et al., 2012]. Such a requirement has been avoided in some previous research, by using internal robot sensors instead. For instance, a method to estimate contact forces from joint torques was presented in [Linderoth et al., 2013]. Further, force controlled assembly without a force sensor was achieved in [Stolt et al., 2012], by estimating contact forces from position errors in the internal controller of the robot.

1.1 Problem formulation

In this work, we address the question of whether robot joint torques could be used to recognize contact force transients during robotic assembly, despite uncertainties introduced by, *e.g.*, joint friction. Further, we investigate how long parts of the

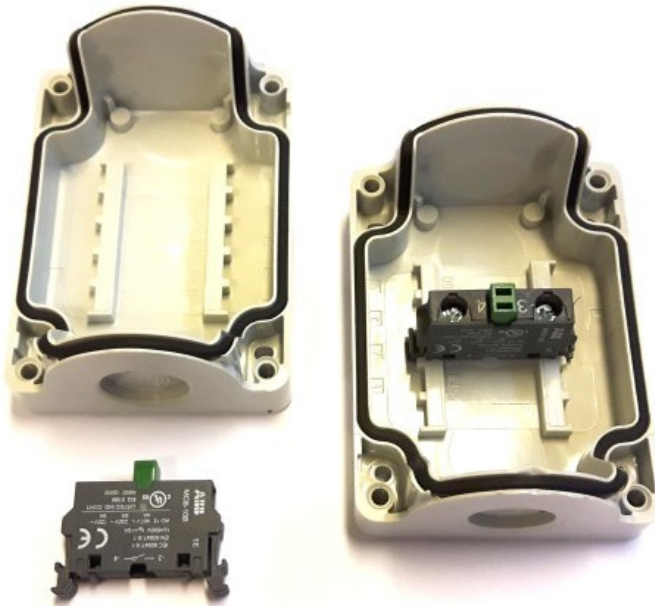


Figure 1. The parts to be assembled. Box (upper left), switch (lower left), and switch attached to box (right).

transients that should be included as input for the detection algorithm, in order to distinguish whether a transient is present or not.

2. Method

The snap-fit scenario considered here consisted of attaching a switch to a box, see Fig. 1. The objective of the robot was to move toward the box while holding the switch, thus pushing the switch against the box, until it snapped into place. The robot should detect the snap-fit automatically, stop moving toward the box, and possibly start a new movement.

2.1 Sequence model

An RNN [Goodfellow et al., 2016; Graves, 2012] was used as a sequence classifier. This choice is discussed in Sec. 5. It had a sequence of joint torques as input, one single output indicating whether the sequence contained a snap-fit or not, one hidden layer, and recurrent connections between its hidden neurons. Each input torque sequence consisted of $T = n_{\text{pre}} + 1 + n_{\text{post}}$ time samples, where n_{pre} and n_{post} were determined as explained in Sec. 2.3. In turn, each time sample consisted of $n_{\text{ch}} = 7$

channels; one per robot joint. The dimension of the hidden layer was chosen to be the same as the number of input channels, n_{ch} .

Denote by $h^{(t)}$ the activation of the hidden units at time step t . The activation was defined recursively as

$$h^{(1)} = \tanh(b + Ux^{(1)}) \quad (4.1)$$

$$h^{(t)} = \tanh(b + Wh^{(t-1)} + Ux^{(t)}) \quad t \in [2; T] \quad (4.2)$$

where U and W are weight matrices, both of size $n_{\text{ch}} \times n_{\text{ch}}$, b is a bias vector with dimension n_{ch} , and $x^{(t)}$ is the input at time t . Further, $\tanh(\cdot)$ represents the hyperbolic tangent function. After reading an entire input sequence, the RNN produced one output $o^{(T)}$ given by

$$o^{(T)} = c + Vh^{(T)} \quad (4.3)$$

where V is a weight matrix of size $2 \times n_{\text{ch}}$, and c is a bias vector with dimension 2. Finally, the softmax operation was applied to generate \hat{y} , a vector that represented the normalized probabilities of the output elements.

$$\hat{y} = \left[\frac{e^{o_1^{(T)}}}{e^{o_1^{(T)}} + e^{o_2^{(T)}}} \quad \frac{e^{o_2^{(T)}}}{e^{o_1^{(T)}} + e^{o_2^{(T)}}} \right]^T \quad (4.4)$$

Here, $o_i^{(T)}$ represents the i :th element of the output vector. If the first element of \hat{y} was larger than the second, or equivalently, larger than 0.5, the data point was classified as positive, *i.e.*, it was indicated that a snap-fit occurred within the sequence. Vice versa, if the first element was less than 0.5, the data point was classified as negative, meaning that no snap-fit was present. The RNN architecture is visualized in Fig. 2.

2.2 Gathering of training data and test data

Training data and test data were obtained as follows. The right arm of the robot was used to grasp the switch, just above the box, as shown in Fig. 3. Thereafter, a reference velocity was sent to the internal controller of the robot, causing the robot gripper to move toward the box at 1.5 mm/s, thus pushing the switch against the box. Once the switch was snapped into place, the robot was stopped manually by the robot operator. The robot joint torques were recorded in 250 Hz through the ABB research interface EGMRI. The torque transient, induced by the snap-fit, was labeled manually, and used to form a positive data point. This procedure was repeated $N = 50$ times, which yielded 50 positive data points. Data prior to each transient was used to form negative data points.

Given n_{pre} and n_{post} , a positive data point was formed by extracting a torque sequence, from n_{pre} samples previous to the peak value of the transient (inclusive), to n_{post} samples after (inclusive). Negative data points, with the same sequence length T as the positive ones, were extracted from torque measurements that ranged from

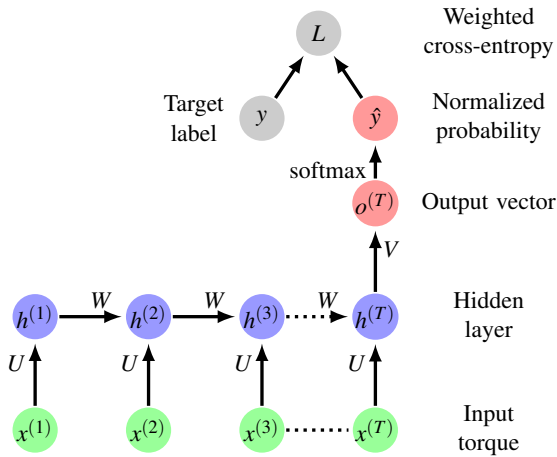


Figure 2. The RNN visualized as an unfolded computational graph, where each node is associated with a certain time step. The biases b and c , as well as the activation function $\tanh(\cdot)$, are omitted for a clearer view, but the computations are detailed in (4.1) to (4.4). The input torque was used to determine the hidden state, which was updated each time step. The last hidden state was used to determine the normalized probability \hat{y} of whether a snap-fit was present in the time sequence or not.

a couple of seconds before the snap-fit, until the positive data point (exclusive). The negative data points were chosen so that overlap was avoided. For each data point, the target was labeled as a two-dimensional one-hot vector y , where $y = [1 \ 0]^T$ represented a positive data point, and $y = [0 \ 1]^T$ represented a negative one.

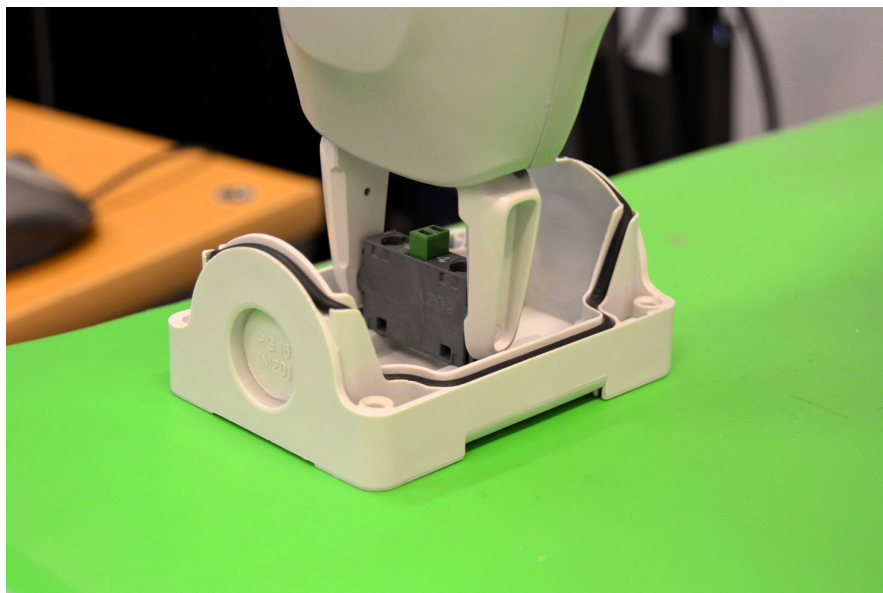
Note that with the approach above, it was possible to extract several negative data points, but only one positive data point, for every snap-fit experienced by the robot.

Half of the positive and negative data points were used in the training set, and the other half was used in the test set.

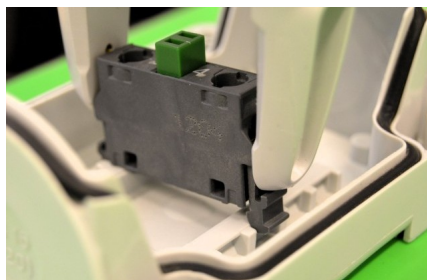
2.3 Model training

Given the training set, the model parameters U, V, W, b , and c were determined by minimizing a loss function L .

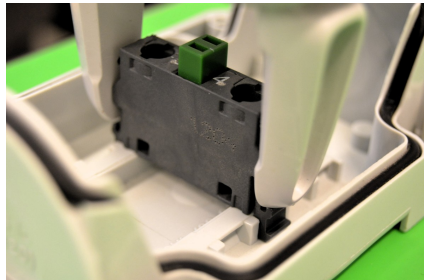
The training set contained much more negative data points than positive ones. If not taken into account, this type of class imbalance has been reported to obstruct the training procedure of several different classifiers. The phenomenon has been described in more detail in [Japkowicz and Stephen, 2002; Japkowicz, 2000], and should be taken into account when designing the loss function. Consider first the following loss function \bar{L} , which is the ordinary cross-entropy between training data



(a)



(b)



(c)

Figure 3. Experimental setup. An overview is shown in (a). The ABB YuMi [ABB Robotics, 2017] robot was used to grasp the switch, and attaching it to the box by pushing downwards. The downward motion began in (b), where the switch was not yet snapped into place. It ended when the snap-fit assembly was complete (c). These photos were taken during the experimental evaluation (see Sec. 3), and the same setup was used for gathering training and test data (see Sec. 2.2).

and model predictions, averaged over the training examples.

$$\bar{L} = -\frac{1}{D} \sum_{d=1}^D \sum_{a=1}^A y_a^d \log \hat{y}_a^d \quad (4.5)$$

Here, a and d are indices for summing over the vector elements and training data points, respectively. This cross-entropy is commonly used as a loss function in machine learning [Goodfellow et al., 2016; Rubinstein and Kroese, 2013]. Due to the class imbalance in the present training set, it would be possible to yield a relatively low loss \bar{L} by simply classifying all or most of the data points as negative, regardless of the input, even though that strategy would not be desirable.

In order to take the class imbalance into account, weighted cross-entropy was used as loss function. Denote by r the ratio between negative and positive data points in the training set, and introduce the weight vector $w_r = [r \ 1]^T$. The loss function was defined as

$$L = -\frac{1}{D} \sum_{d=1}^D \sum_{a=1}^A y_a^d \log \hat{y}_a^d \cdot w_r^T y^d \quad (4.6)$$

The RNN in Sec. 2.1 was implemented as a computational graph in the Julia programming language [Bezanson et al., 2014], using TensorFlow [Abadi et al., 2016; TensorFlow, 2017] and the wrapper TensorFlow.jl [J. Malmaud, 2017]. The Adam algorithm [Kingma and Ba, 2014] was used for minimization of the loss function L .

The values of n_{pre} and n_{post} were determined using both the training set and the test set as follows. All positive data points available were used, and $r = 20$ times as many negative data points. Starting with $n_{\text{pre}} = n_{\text{post}} = 1$, the model was trained using the training set, and its performance was measured using the test set. Subsequently, both n_{pre} and n_{post} were increased by 1, and the training and evaluation procedure was repeated. This continued until perfect classification was achieved, or until the values of n_{pre} and n_{post} were large. (30 was chosen as an upper limit, though it was never reached in the experiments presented here.) Thereafter, n_{pre} was kept constant, and it was investigated how much n_{post} could be lowered with retained performance. This was done by decreasing n_{post} one step at a time, while repeating the training and evaluation procedure for each value. Once the performance was decreased, the value just above that was chosen for n_{post} . This way, the lowest possible value of n_{post} was found, that resulted in retained performance.

Once n_{pre} and n_{post} were determined, new model parameters were obtained by training on a larger data set, with $r = 100$. The reason for using a lower value for the other iterations, was that it took significantly longer computation time to use such a large data set.

Due to the class imbalance in the test set, ordinary classification accuracy, as defined by the number of correctly classified test data points divided by the total number of test data points, would not be a good model performance measurement.

Instead, the F-measurement [Hripcsak and Rothschild, 2005] was used, defined as

$$F_1 = 2 \frac{PR}{P+R} \quad (4.7)$$

where P is the precision, *i.e.*, the number of correctly classified positive data points divided by the number of all data points classified as positive by the model, and R is the recall, *i.e.*, the number of correctly classified positive data points divided by the number of all data points that were truly positive. The value of F_1 ranges from 0 to 1, where 1 indicates perfect classification.

2.4 Implementation of real-time application

The ABB YuMi robot [ABB Robotics, 2017] was used for experimental evaluation. The robot is shown in Fig. 4. The RNN model obtained according to Secs. 2.1 to 2.3 was saved on a server, and loaded into a Julia program on a PC, which communicated with the internal controller of the robot through the LabComm protocol [LabComm, 2017]. The sample frequency was 250 Hz. The robot joint torques were logged and saved, and for each time sample t , a joint torque sequence was formed by the samples in $[t - n_{\text{pre}} - n_{\text{post}}; t]$, and sent as input to the RNN. Measurements after time t were not available at t , which is why the input only contained samples up until t . Each sequence was classified in real-time. The computation time for one classification was short; well below the sample period. To move the robot, desired velocity references for the gripper in Cartesian space were first specified in the Julia program. Then, the corresponding joint velocities were computed using the robot Jacobian, and these were sent as references to the internal controller of the robot.

3. Experiments

Since the test set was used to determine the hyper parameters of the RNN, *i.e.*, n_{pre} and n_{post} , it was necessary to gather new measurements to evaluate the general performance of the RNN. The experimental setup was similar to that in Sec. 2.2, except that the measured torque sequences were saved and classified by the RNN, instead of just saved to the training and test sets. The implementation in Sec. 2.4 was used for robot control and snap-fit detection. The robot was programmed to first move its gripper down, thus pushing the switch against the box. Once a snap-fit was detected, it was programmed to stop its downward motion, and instead move the box to the side. The snap-fit assembly was repeated 50 times, to evaluate the robustness of the proposed approach. The experimental setup is visualized in Fig. 3.

4. Results

The performance of the RNN on the test set, for different values of the hyper parameters, is shown in Table 1. The abbreviations are as follows: number of true

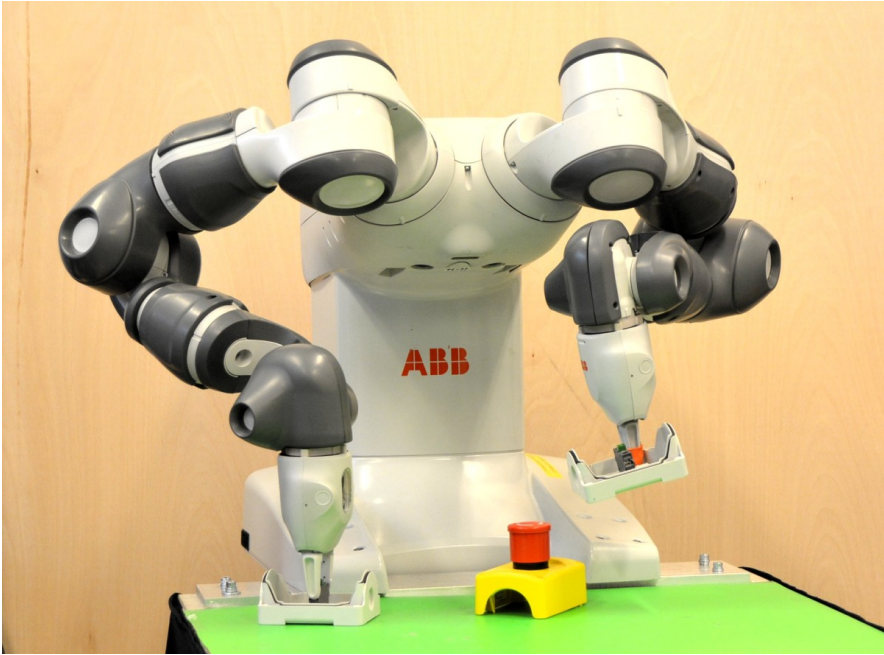


Figure 4. The ABB YuMi robot [ABB Robotics, 2017] used in the experiments.

positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). The hyper parameters were increased until $n_{\text{pre}} = n_{\text{post}} = 5$, for which perfect classification was obtained. Then, n_{post} was decreased until the performance decreased at $n_{\text{post}} = 2$. With this value of n_{post} , larger values of n_{pre} were tested (see second last row in Table 1), which did not yield perfect classification for any values, *i.e.*, $F_1 < 1$. Thus, $(n_{\text{pre}}, n_{\text{post}}) = (5, 3)$ was chosen for the final model.

After training, the RNN detected all 50 snap-fits in the experiments in Sec. 3 correctly, without any false positives prior to the snap. The torque data and RNN output from one of the trials are shown in Figs. 5 and 6. The other trials gave qualitatively similar results.

5. Discussion

There are several alternatives to RNN for classification. Using an RNN is motivated as follows. Two simpler models, matched filter and logistic regression, were tried initially, without achieving satisfactory performance on test data. RNNs are specialized in processing sequential data, and the torque measurements used in this work were sequential. Thanks to the parameter sharing of the RNN, it is possible to estimate a model with significantly fewer training examples, than would be needed

Table 1. RNN performance on the test set, for different values of the hyper parameters. The row with the lowest value of n_{post} that yielded perfect classification is marked in **blue**. With these values, the model was trained and tested again, but now with more negative data points (see last row, marked in **red**).

n_{pre}	n_{post}	TP	TN	FP	FN	P	R	F_1
1	1	20	500	0	5	1	0.80	0.89
2	2	22	500	0	3	1	0.88	0.94
3	3	23	498	2	2	0.92	0.92	0.92
4	4	23	500	0	2	1	0.92	0.96
5	5	25	500	0	0	1	1	1
5	4	25	500	0	0	1	1	1
5	3	25	500	0	0	1	1	1
5	2	22	500	0	3	1	0.88	0.94
[6;30]	2	-	-	-	-	-	-	<1
5	3	25	2500	0	0	1	1	1

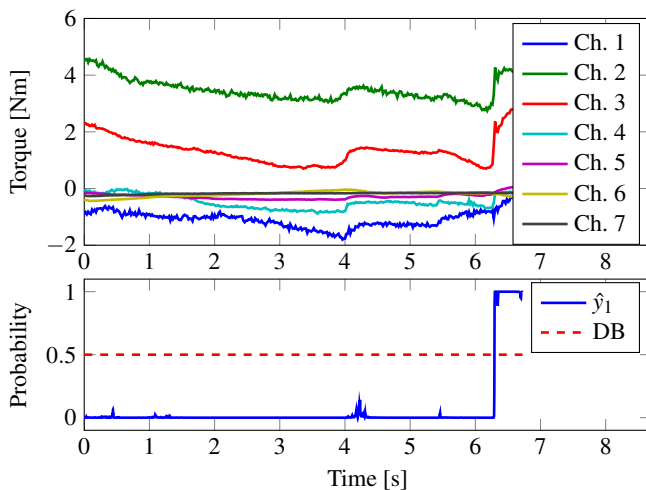


Figure 5. Data from one of the experiments. The robot joint torques (upper plot) were used as input for the RNN. These are represented by one channel (Ch.) per joint. The first element of the RNN output (\hat{y}_1 in the lower plot) was close to 0 before the snap-fit occurred, and increased to close to 1 at the time of the snap-fit. A snap-fit was indicated when \hat{y}_1 was above the decision boundary (DB, at 0.5) for the first time. Thus, for detection purposes, the RNN output generated after this event was not relevant.

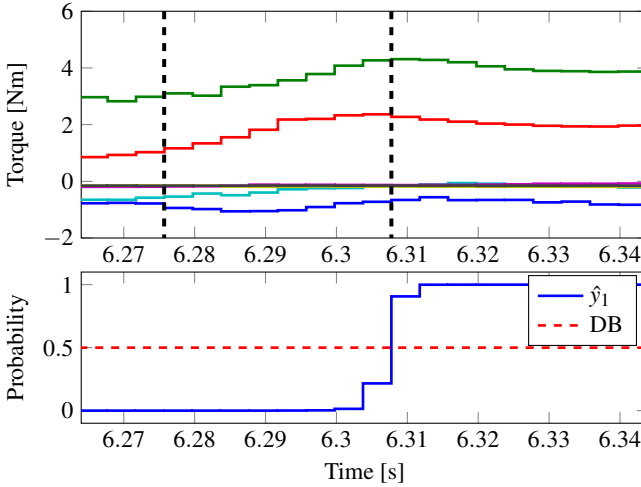


Figure 6. Same data as in Fig. 5, but zoomed-in on the time of the snap-fit. The legend of the upper plot is absent for better visualization, but can be found in Fig. 5. The snap-fit was detected at time $t = 6.308$ s. The first torque sequence to be classified as positive was that within the vertical dashed lines in the upper plot.

without parameter sharing. Compared to models that are not specialized in sequential data, *e.g.*, logit models, SVMs, and ordinary neural networks, the RNN is less sensitive to variations of the exact time step in which some information in the input sequence appears. An RNN can also be generalized to classify data points of sequence lengths not present in the training set, though this was not used in this present work. General properties of RNNs are well described in [Goodfellow et al., 2016].

Compared to [Stolt et al., 2015], the approach proposed here had three new benefits. A force sensor was no longer required, the detection delay was reduced, and the computation time for model training was shortened. In [Stolt et al., 2015], $n_{\text{post}} > 10$ (corresponding to > 40 ms) was required for perfect classification, whereas our proposed method required $n_{\text{post}} = 3$ (12 ms). The training time of the RNN was in the order of minutes on an ordinary PC, which was an improvement compared to days in [Stolt et al., 2015].

The concept of weighted loss to compensate for class imbalance in machine learning has been evaluated in [Japkowicz and Stephen, 2002], and successfully applied to a deep neural network in [Panchapagesan et al., 2016]. Equation (4.6) extends (4.5) by the factor $w_r^T y^d$, which evaluates to r for positive data points, and to 1 for negative ones.

In Fig. 5, it should be noted that \hat{y}_1 raised significantly above 0 at $t \approx 4.2$ s, even though no snap-fit occurred at that time. Even though the values were still

well below the decision boundary, this leaves room for improvement in terms of robustness of the detection.

Given a certain contact force/torque acting on the end-effector of the robot, the corresponding joint torques depend on the configuration, as well as gravity and friction. In order to generalize the detection to other robot states than the one used in the training, it would be a good idea to estimate the contact forces/torques, by first modeling the gravity and friction-induced torques, and subsequently compensating for these.

Whereas the evaluation of the model was done in real-time in this work, the RNN training was performed offline. It therefore remains as future work to create a user interface that allows for an operator to gather training data and test data, label them, and run the training procedure.

It is also important to shorten the reaction time of the robot, *i.e.*, to further reduce the value of n_{post} . This could be done by including more sensors. For instance, the snap-fit assembly generates a sound, easily recognized by a human. Adding a microphone to the current setup, would therefore add information to the detection approach. In turn, this could be used to decrease the required amount of training data, improve robustness of the detection, or detect the snap-fit earlier.

6. Conclusion

In this work, we have addressed the question of whether robot joint torque measurements, together with an RNN model, could be used for detection of snap-fits during assembly. First, training and test data were gathered and labeled. Then, these were used to determine the model parameters. Finally, a real-time application for snap-fit detection was implemented and tested. The method presented seems promising, since the resulting model had high performance, both on the test data and during the experiments.

Acknowledgments

The authors would like to thank Maj Stenmark and Jacek Malec at Computer Science, Lund University, for valuable discussions throughout this work. Fredrik Bagge Carlson at Dept. Automatic Control, Lund University, and Mathias Haage at Computer Science, Lund University, are gratefully acknowledged both for valuable discussions, and for the development of a bridge to the research interface EGMRI. The authors are members of the LCCC Linnaeus Center and the ELLIIT Excellence Center at Lund University. The research leading to these results has received funding from the European Commission's Framework Programme Horizon 2020 – under grant agreement No 644938 – SARAFun.

References

- Abadi, M., A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, et al. (2016). “TensorFlow: large-scale machine learning on heterogeneous distributed systems”. *arXiv preprint arXiv:1603.04467*.
- ABB Robotics (2017). *ABB YuMi*. URL: <http://new.abb.com/products/robotics/yumi> (visited on 2016-01-23).
- Bezanson, J., A. Edelman, S. Karpinski, and V. B. Shah (2014). “Julia: a fresh approach to numerical computing”. *arXiv:1411.1607*. eprint: 1411.1607 (cs.MS).
- Goodfellow, I., Y. Bengio, and A. Courville (2016). *Deep Learning*. Accessed: 2017-03-06. MIT Press, Cambridge, MA. URL: <http://www.deeplearningbook.org>.
- Graves, A. (2012). “Neural networks”. In: *Supervised Sequence Labelling with Recurrent Neural Networks*. Springer, Berlin Heidelberg, Germany, pp. 15–35.
- Hripcsak, G. and A. S. Rothschild (2005). “Agreement, the F-measure, and reliability in information retrieval”. *Journal of the American Medical Informatics Association* **12**:3, pp. 296–298.
- J. Malmaud (2017). *A Julia wrapper for TensorFlow*. URL: <https://github.com/malmaud/TensorFlow.jl> (visited on 2017-03-09).
- Japkowicz, N. (2000). “The class imbalance problem: Significance and strategies”. In: *International Conference on Artificial Intelligence*. June 26–29, Las Vegas, Nevada.
- Japkowicz, N. and S. Stephen (2002). “The class imbalance problem: A systematic study”. *Intelligent Data Analysis* **6**:5, pp. 429–449.
- Kingma, D. and J. Ba (2014). “Adam: a method for stochastic optimization”. *arXiv preprint arXiv:1412.6980*.
- LabComm (2017). *Research tools and software*. Dept. Automatic Control, Lund University. URL: <http://www.control.lth.se/Research/tools.html> (visited on 2017-03-06).
- Linderoth, M., A. Stolt, A. Robertsson, and R. Johansson (2013). “Robotic force estimation using motor torques and modeling of low velocity friction disturbances”. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. November 3–7, Tokyo, Japan, pp. 3550–3556.
- Panchapagesan, S., M. Sun, A. Khare, S. Matsoukas, A. Mandal, B. Hoffmeister, and S. Vitaladevuni (2016). “Multi-task learning and weighted cross-entropy for DNN-based keyword spotting”. *Interspeech 2016*, pp. 760–764.
- Rodriguez, A., D. Bourne, M. Mason, G. F. Rossano, and J. Wang (2010). “Failure detection in assembly: force signature analysis”. In: *IEEE Conference on Automation Science and Engineering (CASE)*. August 21–24, Toronto, Canada, pp. 210–215.

- Rojas, J., K. Harada, H. Onda, N. Yamanobe, E. Yoshida, K. Nagata, and Y. Kawai (2012). “A relative-change-based hierarchical taxonomy for cantilever-snap assembly verification”. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. October 7–12, Vilamoura, Portugal, pp. 356–363.
- Rubinstein, R. Y. and D. P. Kroese (2013). *The Cross-Entropy Method: A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation and Machine Learning*. Springer Science & Business Media, New York.
- Stolt, A., M. Linderoth, A. Robertsson, and R. Johansson (2012). “Force controlled robotic assembly without a force sensor”. In: *IEEE International Conference on Robotics and Automation (ICRA)*. May 14–18, St. Paul, MN, pp. 1538–1543.
- Stolt, A., M. Linderoth, A. Robertsson, and R. Johansson (2015). “Detection of contact force transients in robotic assembly”. In: *IEEE International Conference on Robotics and Automation (ICRA)*. May 26–30, Seattle, WA, pp. 962–968.
- TensorFlow (2017). *An open-source software library for machine intelligence*. URL: <https://www.tensorflow.org/> (visited on 2017-03-09).