

Diss. ETH No. 20593

Improving Tracking Performance by Learning from Past Data

A dissertation submitted to
ETH ZURICH

for the degree of
Doctor of Sciences

presented by

ANGELA P. SCHOELLIG

Diplom-Ingenieurin (Technische Kybernetik), Universität Stuttgart
MSc in Engineering Science and Mechanics, Georgia Institute of Technology, Atlanta

born July 17, 1983
citizen of Germany

accepted on the recommendation of

Prof. Dr. Raffaello D'Andrea, examiner
Prof. Dr. Andrew Alleyne, co-examiner

2012

Abstract

The main theme of this thesis is the development of machine learning algorithms for high-performance trajectory tracking. We consider dynamic systems that are required to precisely follow predefined trajectories. The goal of this research is to explore how past data (for example measurements from previous executions) can be used to improve a system's tracking performance.

A typical means of imposing a desired behavior on a dynamic system is feedback control. In such a setup, the motion of the system is guided by an external reference signal and the influence of noise and unexpected disturbances is reduced by feeding back the measured system output. The design of feedback control systems is often based on a mathematical model of the underlying system. The performance of such control schemes is limited by the accuracy of the dynamics model and the causality of the control action that is compensating for disturbances only as they occur.

We address these limitations by proposing a data-based control approach that is able to store and interpret information from past experiments, and infer the correct control actions for future performances. This research is motivated by recent computational advances, which provide enormous possibilities for storing, processing and evaluating large amounts of data. We aim to exploit these new possibilities with three main contributions:

First, we present an algorithm that exploits data from a repeated operation in order to learn to precisely follow a predefined trajectory. We adapt the feed-forward reference signal to the system with the goal of achieving high tracking performance – even under the presence of model errors and other recurring disturbances. The approach is based on a coarse model of the system dynamics and uses measurements from past executions to optimize the tracking performance. We combine traditional optimal filtering methods with state-of-the-art optimization techniques in order to obtain an effective and computationally efficient learning strategy. The proposed approach falls into the area of iterative learning control. Novel features of our approach are the direct treatment of input and state constraints when updating the feed-forward reference, an identification routine that extracts the required system model from a numerical simulation, and a termination condition that stops an execution early if the deviation from the nominal trajectory exceeds a given bound. The latter allows for a safe learning that gradually extends the time horizon of the trajectory. These new features are particularly relevant when we apply the algorithm to highly maneuverable quadrotor vehicles in the ETH Flying Machine Arena. We aim to exploit their full dynamic potential and to improve on time-optimized trajectories. The learning scheme has proven to be effective both when directly learning the thrust and rotational rate inputs sent to the quadcopter, and when building the learning scheme on

Abstract

top of a quadcopter system that is guided by a trajectory following controller. The numerical identification routine was used in the latter case to avoid extensive analytical modeling.

For the second project, we consider iterative learning control in a multi-agent framework, wherein a group of agents simultaneously and repeatedly perform the same task. Assuming similarity between the agents, we investigate whether exchanging information between the agents improves an individual's learning performance. That is, does an individual agent benefit from the experience of the other agents? We derive analytical bounds on the performance improvement due to joint learning.

The third project uses learning in a different context. Here, we aim to precisely track periodic trajectories with a quadcopter. We develop a learning scheme that determines feed-forward correction parameters for a large class of periodic motions from a small set of identification experiments. This research is motivated by our vision of designing and executing multi-vehicle flight that is coordinated to music.

In addition to these three main results, we studied the dynamic limits of quadrotor vehicles and developed algorithms for both generating feasible trajectories, and checking the feasibility of a given trajectory. The first is important for the iterative learning project, where the goal is to perform and improve on highly agile motions. We must therefore design trajectories that are dynamically challenging but still feasible. The latter is important for the design of rhythmic performances, where we wish to check the feasibility of a choreography prior to actual flight.

Aerial robots and, in particular, quadcopters are a great platform for showcasing the algorithms we have developed to both scientific and non-scientific audiences. Several demonstrations were developed with the goal of visually communicating the key concepts of this work to a large audience.

Zusammenfassung

Im Mittelpunkt dieser Dissertation steht die Entwicklung von Algorithmen für maschinelles Lernen. Das Ziel der Algorithmen ist es, hochgenaue Trajektorienfolge zu erreichen. Wir betrachten dynamische Systeme, die präzise einer vorgegebenen Trajektorie folgen sollen, und untersuchen, inwiefern Daten aus vorherigen Experimenten (zum Beispiel Messwerte von früheren Durchführungen eines Experiments) verwendet werden können, um die Folgegenauigkeit eines Systems zu verbessern.

Eine klassische Methode, um ein gewünschtes Systemverhalten zu erreichen, ist die Regelung mit Hilfe einer Ausgangsrückführung. Dabei wird das dynamische Verhalten eines Systems durch ein externes Stellsignal gesteuert. Der Einfluss von Rauschen und unerwarteten Störungen wird durch die Rückführung des Systemausgangs reduziert. Dieser Regelkreis wird oftmals basierend auf einem mathematischen Modell des zugrunde liegenden Systems entworfen. Die Güte einer solchen Regelstrategie ist beschränkt durch die Genauigkeit des dynamischen Modells und durch die Kausalität des Regelvorgangs, der auf Störungen erst reagiert, sobald sie auftreten.

Wir stellen einen datenbasierten Regelansatz vor, der die genannten Probleme aufgreift und löst. Informationen von vorhergegangenen Experimenten werden gespeichert und gedeutet. Aus diesen Daten werden die richtigen Regelstrategien für zukünftige Experimente abgeleitet. Jüngste technologische Fortschritte, die es ermöglicht haben, große Datenmengen zu speichern, zu verarbeiten und auszuwerten, machen einen solchen Ansatz möglich. Unser Ziel ist es, diese neuen technologischen Möglichkeiten auszuschöpfen. In dieser Dissertation machen wir dazu drei wesentliche Beiträge:

Das erste Ergebnis ist ein Lernalgorithmus, der Messdaten von dem wiederholten Durchführen eines Experiments nutzt, um zu lernen präzise einer vorgegebenen Trajektorie zu folgen. Nach jedem Durchgang wird das Steuersignal (das heißt, das externe Stellsignal; auch Führungsgröße genannt) angepasst, mit dem Ziel hohe Folgegenauigkeit trotz Modellfehlern und anderen wiederkehrenden Störungen zu erreichen. Der Ansatz basiert auf einem einfachen mathematischen Modell des dynamischen Systems und verwendet Messdaten von früheren Durchgängen, um die Folgegenauigkeit zu verbessern. Der Algorithmus verbindet klassische Methoden der optimalen Filterung mit modernsten Optimierungsverfahren. Das Ergebnis ist eine effektive und recheneffiziente Lernstrategie. Der entwickelte Lernalgorithmus fällt in die Kategorie der Iterativ-Lernenden-Regelung (engl.: Iterative Learning Control). Wesentliche Neuerungen unseres Ansatzes sind die direkte Berücksichtigung von Eingangs- und Zustandsgrößen-Beschränkungen bei der Berechnung des neuen Steuersignals, eine Identifikationsroutine, die das benötigte Systemmodell aus einer numerischen Simulation ableitet, und

Zusammenfassung

eine Abbruchbedingung, die einen Durchgang frühzeitig stoppt, falls die Abweichung von der nominellen Trajektorie eine gegebene Grenze übersteigt. Letzteres ermöglicht ein gefahrloses Lernen, bei dem der Zeithorizont der Trajektorie schrittweise verlängert wird. Die Besonderheiten des neuen Algorithmus spielen insbesondere für unsere Anwendung eine wichtige Rolle. Wir evaluieren den Ansatz auf kleinen, wendigen Quadroptern in der ETH Flying Machine Arena. Unser Ziel ist dabei das dynamische Potenzial der Quadropten auszunutzen und zeitoptimierte Trajektorien zu lernen. Das Lernverfahren hat sich als erfolgreich erwiesen, sowohl wenn die Eingangsgrößen des Quadropters (Gesamt-Schub und Drehraten) direkt gelernt werden, als auch wenn der Lernalgorithmus auf ein Quadropter-System angewendet wird, das von einem Trajektorien-Folgeregler gesteuert wird. Für Letzteres wurde die erwähnte numerische Identifikationsroutine verwendet, um das aufwändige Herleiten eines analytischen Modells zu umgehen.

Das zweite Projekt beschäftigt sich mit der Iterativ-Lernenden-Regelung für Multiagenten-Systeme. Genauer gesagt betrachten wir eine Gruppe von Agenten, die gleichzeitig und wiederholt die gleiche Aufgabe ausführen. Unter der Annahme, dass eine Ähnlichkeit zwischen den Agenten besteht, untersuchen wir, ob der Austausch von Information zwischen den Agenten den Lernerfolg eines einzelnen Agenten erhöht. Das heißt: Profitiert ein einzelner Agent von den Erfahrungen der anderen Agenten? Wir leiten eine analytische Schranke her, die die Verbesserung des Lernerfolgs durch ein gemeinschaftliches Lernen nach oben beschränkt.

Das dritte Projekt betrachtet Lernen in einem anderen Zusammenhang. Das Ziel ist hier, dass Quadropten periodischen Trajektorien präzise folgen. Wir entwickeln einen Lernalgorithmus, der Steuersignal-Korrekturparameter für eine große Klasse periodischer Bewegungen mithilfe einer sehr begrenzten Anzahl von Identifikationsexperimenten bestimmt. Motivation für diese Arbeit ist unsere Vision, Flugdarbietungen für mehrere Quadropten abgestimmt zur Musik zu entwerfen und auszuführen.

Neben diesen drei Hauptresultaten, haben wir insbesondere die dynamischen Grenzen von Quadropten untersucht und Algorithmen entwickelt, die (für Quadropten) ausführbare Trajektorien generieren und die die Ausführbarkeit einer gegebenen Trajektorie testen. Ersteres ist wichtig für das iterative Lernen. Hier ist unser Ziel außerordentlich schnelle Bewegungen zu lernen. Wir müssen daher Trajektorien generieren, die dynamisch anspruchsvoll, aber dennoch ausführbar sind. Letzteres ist wichtig für das Design von rhythmischen Flug-Darbietungen. Hier ist es entscheidend, die Ausführbarkeit einer Choreographie vor dem eigentlichen Flug zu testen.

Flugroboter und insbesondere Quadropten sind eine großartige Plattform, um die Algorithmen, die wir entwickelt haben, sowohl einem wissenschaftlichen Publikum als auch der breiten Öffentlichkeit zu präsentieren. Verschiedene Vorführungen wurden entwickelt, mit dem Ziel visuell die grundlegenden Konzepte dieser Arbeit einem großen Publikum zu vermitteln.