Improving tracking performance by learning from past data

Angela P. Schoellig
Doctoral examination – July 30, 2012

Advisor: Prof. Raffaello D’Andrea // Co-advisor: Prof. Andrew Alleyne
Improving tracking performance by learning from past data = experience

Angela P. Schoellig
Doctoral examination – July 30, 2012

Advisor: Prof. Raffaello D’Andrea // Co-advisor: Prof. Andrew Alleyne
MOTIVATION

HUMANS learn from experience.

We learn from mistakes and get better through practice.

We constantly adapt to changing environments.
MOTIVATION

AUTOMATED SYSTEMS typically make the *same mistakes* over and over again when performing a task repeatedly. *Why?*

Robots of a car assembly line.
AUTOMATED SYSTEMS are typically operated using feedback control:

Performance limitations:
- Causality of disturbance correction: “first detect error, then react”.
- Model-based controller design; model ≠ real system.
GOAL

Improve the performance over causal, feedback control by learning from previous experiments.
SCOPE OF WORK

Learning task:
Following a predefined trajectory.

Approach:
• Model-based learning based on *a priori* knowledge of the system dynamics.
• Adaptation of the input.

Potential:
Acausal action, anticipating repetitive disturbances.
OVERVIEW

I. Introduction
   a. Testbed: The Flying Machine Arena
   b. Motivation for learning

II. Project A. Iterative learning for precise trajectory following: single-agent and multi-agent results. Focus of this talk

III. Project B. Learning of feed-forward parameters for rhythmic flight performances

IV. Summary
TESTBED, see [www.flyingmachinearena.org](http://www.flyingmachinearena.org)
THE TEAM

Mark Müller

Markus Hehn

Sergei Lupashin

Federico Augugliaro
THE FLYING MACHINE ARENA

Vehicle position and attitude

Control Algorithms

Collective thrust and turn rates (wireless)
OPERATION

Input
Desired position

Trajectory-following controller (TFC)

Collective thrust and turn rates

Output
Measured position and attitude
MOTIVATION: PROJECT A

Desired motion.
MOTIVATION: PROJECT A

Performance with trajectory-following controller.

Different trials

Large repetitive error
OVERVIEW

I. Introduction

II. Project A. Iterative learning for precise trajectory following
   a. Learning approach
   b. Results

III. Project B. Learning of feed-forward parameters for rhythmic flight performances

IV. Summary
Peer-reviewed publications


Joint work with Fabian L. Mueller (Master student).
**A | LEARNING APPROACH**

**Features:** Learning through a repeated operation, updating full input trajectory after each trial.

---

**Diagram:**
- INPUT trajectory → SYSTEM → OUTPUT trajectory
- Updated input → INPUT UPDATE → DISTURBANCE ESTIMATION → Updated disturbance
- Do it again!
PREREQUISITES

- **Dynamics model of system**
  - (i) in analytical form
  - (ii) in form of a numerical dynamics simulation

- **Desired output trajectory** $y^*(t), t \in [0, t_f]$, and corresponding nominal input trajectory $u^*(t)$.
  - $(u^*(t), y^*(t))$ must satisfy the model equations.

RESULT

- Learned input
- Estimated disturbance vector
Dynamics model of the physical system: \( \dot{x}(t) = f(\ddot{x}(t), \ddot{u}(t)), \quad \ddot{y}(t) = \ddot{x}(t). \)

Consider small deviations from nominal trajectory.

\[
\tilde{u}(t) = \ddot{u}(t) - u^*(t), \quad \tilde{x}(t) = \ddot{x}(t) - x^*(t), \quad \tilde{y}(t) = \ddot{y}(t) - y^*(t)
\]

Linearize and discretize. Linear, time-varying difference equation.

\[
\tilde{x}(k+1) = A_D(k)\tilde{x}(k) + B_D(k)\tilde{u}(k), \quad \tilde{y}(k) = \tilde{x}(k), \quad k \in \{0, \ldots, N\}.
\]

Static mapping. Representing one trial.

\[
\begin{bmatrix}
\tilde{x}(0) \\
\tilde{x}(1) \\
\tilde{x}(2) \\
\vdots \\
\tilde{x}(N)
\end{bmatrix} =
\begin{bmatrix}
0 & 0 & \cdots & 0 & 0 \\
B_D(0) & 0 & \cdots & 0 & 0 \\
\Phi_{(1,1)}B_D(0) & B_D(1) & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
\Phi_{(N-1,1)}B_D(0) & \Phi_{(N-1,2)}B_D(1) & \cdots & B_D(N) & 0
\end{bmatrix}
\begin{bmatrix}
\tilde{u}(0) \\
\tilde{u}(1) \\
\tilde{u}(2) \\
\vdots \\
\tilde{u}(N)
\end{bmatrix}
\]

With \( \Phi_{(l,m)} = A_D(l)A_D(l+1)\cdots A_D(m), \ l < m, \) and \( \tilde{x}(0) = 0. \)
A | ITERATION-DOMAIN MODEL

For each trial \( j, \ j \in \{1, 2, \ldots \} \),

\[
y_j = F u_j + d_j + \mu_j.
\]

Recurring disturbance \( d_j \).
Unknown. Only small changes between iterations:

\[
d_j = d_{j-1} + \omega_{j-1}.
\]

Noise \( \mu_j \).
Unknown. Changing from iteration to iteration.

\[\mu_j, \omega_j \sim \text{trial-uncorrelated, zero-mean Gaussian noise}\]

From trial to trial our knowledge about \( d_j \) improves.
A | **STEP 1: ESTIMATION**

UPDATE OF DISTURBANCE ESTIMATE via **Kalman filter** in the iteration domain:

estimates the repetitive disturbance \( d_j \) by taking into account all past measurements.

Prediction step:

\[
d_j = d_{j-1} + \omega_{j-1}.
\]

Measurement update step:

\[
y_j = F u_j + d_j + \mu_j.
\]

Obtain \( \hat{d}_{j|j} \).
INPUT UPDATE via **convex optimization:**

minimizes the tracking error in the next trial:

\[ E[\hat{y}_{j+1} | \text{all past measurements}] = F u_{j+1} + \hat{d}_{j|j}. \]

\[
\min_{u_{j+1}} \left\| F u_{j+1} + \hat{d}_{j|j} \right\|_p \\
p \in \{1, 2, \infty\}
\]

subject to

\[ u_{\min} \leq u_{j+1} \leq u_{\max} \]
\[ x_{\min} \leq x_{j+1} \leq x_{\max} \]

**Obtain** \(u_{j+1}\).
## TWO EXPERIMENTAL SCENARIOS

<table>
<thead>
<tr>
<th>SCENARIO 1</th>
<th>SCENARIO 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>• No feedback from motion capture cameras during task execution</td>
<td>• Camera information is used.</td>
</tr>
<tr>
<td>• Analytical model</td>
<td>• Model via numerical simulation</td>
</tr>
<tr>
<td>• 2D quadrocopter model</td>
<td>• 3D quadrocopter model</td>
</tr>
<tr>
<td>• Constraints on single motor thrusts and turn rates.</td>
<td></td>
</tr>
</tbody>
</table>

![SCENARIO 1 Diagram](image1)

![SCENARIO 2 Diagram](image2)
A | SCENARIO 1: state trajectories

S-shaped trajectory.
S-shaped trajectory.
A | SCENARIO 1: state trajectories

S-shaped trajectory.

![Graph showing state trajectories with labels It 0, It 1, It 2, It 3, It 4, It 5, z [m], y [m], and legend indicating blue for Desired, black for Learned It 0–2, and red for Learned It 3–9.]
## TWO EXPERIMENTAL SCENARIOS

### SCENARIO 1
- No feedback from motion capture cameras during task execution
  - Collective thrust and turn rates
  - Position, attitude

### SCENARIO 2
- Camera information is used.
  - Position
  - TFC
  - Position, attitude

- Analytical model
- Model via numerical simulation
- 2D quadrocopter model
- 3D quadrocopter model
- Constraints on single motor thrusts and turn rates.
S-shaped trajectory.

A | SCENARIO 2: state trajectories
A | SCENARIO 2: state trajectories

S-shaped trajectory.
A | SCENARIO 2: state trajectories

S-shaped trajectory.
A | SCENARIO 2: state trajectories

S-shaped trajectory.
A | SCENARIO 2: error convergence
• **Prerequisites:** approximate model of system dynamics.
• **Efficient learning algorithm:** convergence in around 5-10 iterations.
• **Acausal compensation:** outperforms pure feedback control.

**Powerful combination**  Learning applied to feedback-control systems: compensation for repetitive and non-repetitive disturbances.
VIDEO: http://tiny.cc/SlalomLearning

Quadrocopter Slalom Learning
OVERVIEW

I. Introduction

II. Project A. Iterative learning for precise trajectory following

III. Project B. Learning of feed-forward parameters for rhythmic flight performances
   a. Learning approach
   b. Results

I. Summary
B | PUBLICATIONS

Peer-reviewed publications


Joint work with Federico Augugliaro (Bachelor/Master student) and Clemens Wiltsche (semester project).
VIDEO: http://tiny.cc/DanceWith3

Dancing Quadrocopters

Rise Up

IDSC

ETH Zürich
Task: Precise tracking of *periodic* motions.

Features:

- Learning through a dedicated identification routine performed prior to flight performance.
- Adaptation of only a few *input parameters*. 
B | LEARNING APPROACH

PURE FEEDBACK

WITH LEARNED CORRECTION FACTORS

Amplitude and phase error

For each directional motion component and frequency, we learn:
(1) amplitude correction factor,
(2) additive phase correction.
VIDEO: http://tiny.cc/Armageddon
OVERVIEW

I. Introduction

II. Project A. Iterative learning for precise trajectory following

III. Project B. Learning of feed-forward parameters for rhythmic flight performances

IV. Summary
Repetitive error components can be effectively compensated for by learning from past data. Result is an improved tracking performance.
RESEARCH SUPPORT STAFF

Igor Thommen

Carolina Flores

Hans Ulrich Honegger

Marc Corzillius

Thank you!
IT FOLLOWS...

Live demonstration in the Flying Machine Arena
Improving tracking performance by learning from past data

Angela P. Schoellig
Doctoral examination – July 30, 2012

Advisor: Prof. Raffaello D’Andrea // Co-advisor: Prof. Andrew Alleyne