Advanced Robot Control and Learning

Prof. Sami Haddadin

Advanced Robot Control - Outline

Differential Geometry in Robotics
Basics of Task Space Modeling and Control
Redundant Robots
Modern Methods of Robot Control
Port-Hamiltonian Systems in Robotics
Linear Parametrization and Identification of Robot Dynamics
Bio-Inspired Robot Control

- Bio-Inspired Formulation
- Adaptive Impedance Control for a Manipulator

Chapter 7: Bio-Inspired Robot Control

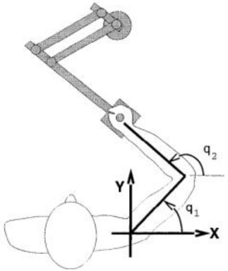
Bio-Inspired Formulation

Motivation

- Unlike animals, humans take a long while to learn to control even basic movements of their body
- But eventually, we learn to do sophisticated movements (e.g. dealing with novel environments such as steep ski slopes!)

But, how can human adapt to these environments?
→ Experimental research on *Neuromechanics and Motor Control*

Task (Shadmehr et al 1994):
Reaching movements in presence of external forces imposed from mechanical environment

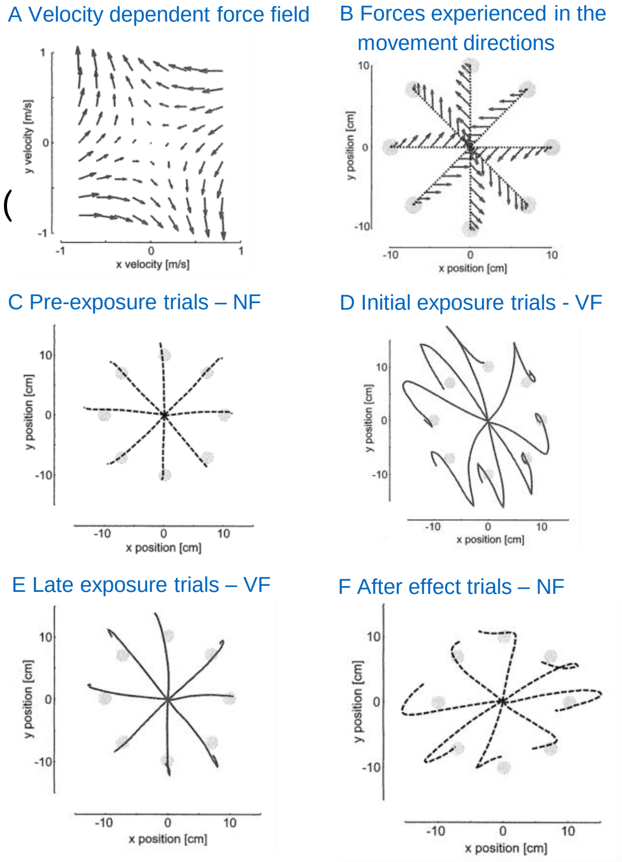


Adaption to Novel Dynamics

Experiments:

1. Point-to-point (PTP) movements on a horizontal plane: Null force field (
   (figure C: Easy!)
2. PTP movements on a horizontal plane under a stable force field:
   e.g. Velocity-dependent force field (*VF*) (figure D: first trial, figure E: later trials)
3. PTP movements on a horizontal plane under VF,
   then suddenly NF (What would happen?)

Figure F: big trajectory disturbances in the *opposite* direction
→ subjects learned to compensate for the forces!
→ controller did not just simply stiffen the joints[1] but
the learned compensation has *feedforward* characteristics



[1] otherwise straighter movements would occur

Shadmehr et al 1994

Adaption to Novel Dynamics

Experiment: Point-to-point movements on horizontal plane under different VF's
(Kinematic error not systematically reduced!)
→ Controller compensates for the approximate mean of the random distribution of the force fields,
→ primarily by the strength of the force field on the most recent trials

Suggested *feedforward* output of the adaptation process[2]:

$$z^{k+1} = \vartheta z^k + \alpha e^k, \qquad \vartheta, \alpha > 0$$

where e^k is the last trial error and z^k is the previously learned output

Consequently, as learning progresses and error is reduced, the feedback sensory signals appear to shift forward in time!

[2] Known as *autoregressive* adaption model

Motor Learning under Unstable and Unpredictable Conditions

Adaptation to novel stable dynamics ↔ Learning the appropriate feedforward forces for each state

Are all novel environments stable?

- Opening a door with high resistance vs. Using a screwdriver (*inherently unstable!*)

Is instability the only problem? NO!

- Unpredictability
- Unrepeatability

How about noise? (both in perception and action)

- Error in state of the body
- Error in state of the world
- Leading to wrong feedforward command

Noise in opening a door with high resistance vs. in using a screwdriver

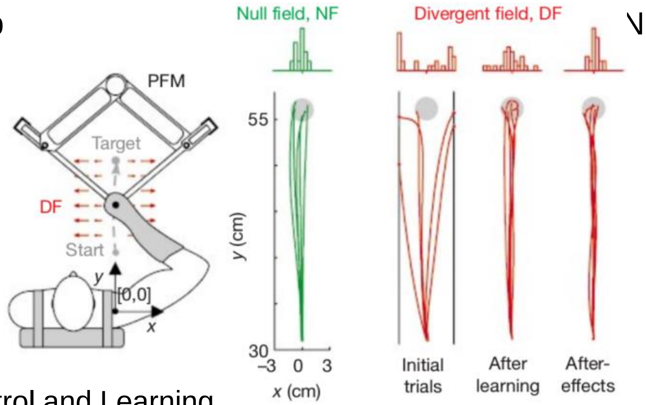
Unstable Interaction

Experiment:

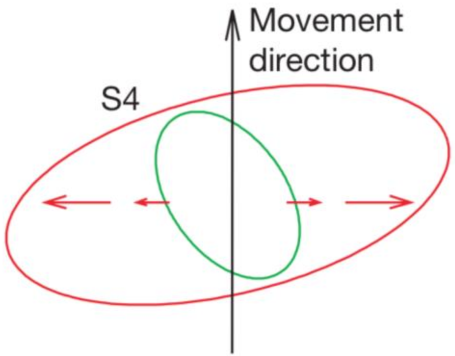
PTP movements on a horizontal plane under an unstable position-dependent divergent force field (DF) of $F_x = K_x x$ and $F_y = 0$, where x is the perpendicular distance to the PTP straight line, F_x and F_y are the forces applied respectively perpendicular and in parallel to the straight line, and K_x is the field constant. (What would happen?)

- Co-contraction will increase the stiffness and viscosity of the joints and consequently the endpoint of limb
- Posture modification will changes the effective mass of the limb and the postural contributions to endpoint stiffness and viscosity

Eventually we can follow a straight line, b ... NF:



Burdet et al. *Nature* (2001)



Adapted Impedance

Limb’s endpoint impedance is tuned:

- To exactly compensate for the imposed instability → *Scalability*
- In direction of instability → *Directionality*

Limb’s endpoint impedance tuning can be done by:

- Co-contraction: Fast but high metabolic cost
  (e.g. when someone hits you and your eyes are open)
- Feedback tuning: Slow but low metabolic cost
  (e.g. when someone hits you and your eyes are closed)

Change of Impedance and Force During Learning

- In both stable and unstable novel environments → An early co-contraction[3] in learning
- Unstable force field → Co-contraction decays in a selective manner to final level of sufficient endpoint stiffness
- Stable force field → After compensation for position error, co-contraction will turn to a feedforward torque, leaving the task to the directly responsible muscles

⇒ Eventual decrease in energy consumption for both cases

[3] known as wasted contraction

MIRMI TUM

Principles of Motor Adaption

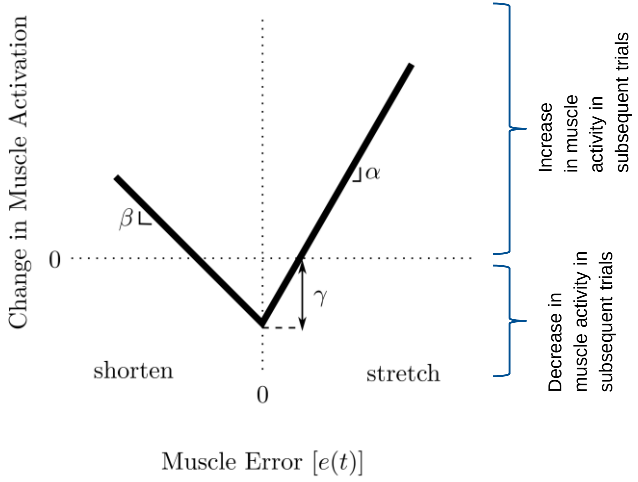
- Motor commands are composed of *feedforward* and *feedback* commands

<u>Principles of motor adaptation:</u>

- ➢ Errors *stretching* a muscle along desired trajectory
  $\rightarrow$ increment in *feedforward* activity,
  proportional to the size of the error (α)
- ➢ Errors *shortening* a muscle along desired trajectory
  $\rightarrow$ increment in *feedforward* activity,
  proportional to the size of the error
  (with lower proportionality constant) (β)
- ➢ Muscle activation is reduced as error decrease with practice
  $\rightarrow$ γ determines vertical offset, acts to minimize activation

V-shaped learning function

(learning parameters $\alpha > \beta$ determine the slope)



Error-Based Learning Algorithm

CNS Learns Stable, Accurate, and Efficient Movements Using a Simple Algorithm, David W. Franklin et al.,Journal of Neuroscience, 2008

- Assumption: each motor command $\boldsymbol{w} \equiv (w_1, w_2, \dots, w_M)$ for m muscels

$$\boldsymbol{w} = \boldsymbol{u} + \boldsymbol{v}$$

with:

- $\boldsymbol{u}$: feedforward term, corresponding to learned dynamics
- $\boldsymbol{v}$: feedback term, corresponding to feedback responses from unlearned dynamics

- Feedforward motor command u_i^k for each muscle i is updated from trial k to $k+1$:

$$\Delta u_i^k(t) = \alpha\left[\varepsilon_i^k(t)\right]_+ + \beta\left[-\varepsilon_i^k(t)\right]_+ - \gamma, \qquad \alpha > \beta > 0, \quad \gamma > 0$$

$$u_i^{k+1}(t) = \left[u_i^k(t) + \Delta u_i^k(t+\zeta)\right]_+, \qquad [.]_+ = \max\{., 0\}$$

$$\varepsilon_i^k(t) = e_i^k(t) + \delta \dot{e}_i^k(t), \qquad \delta > 0$$

with:

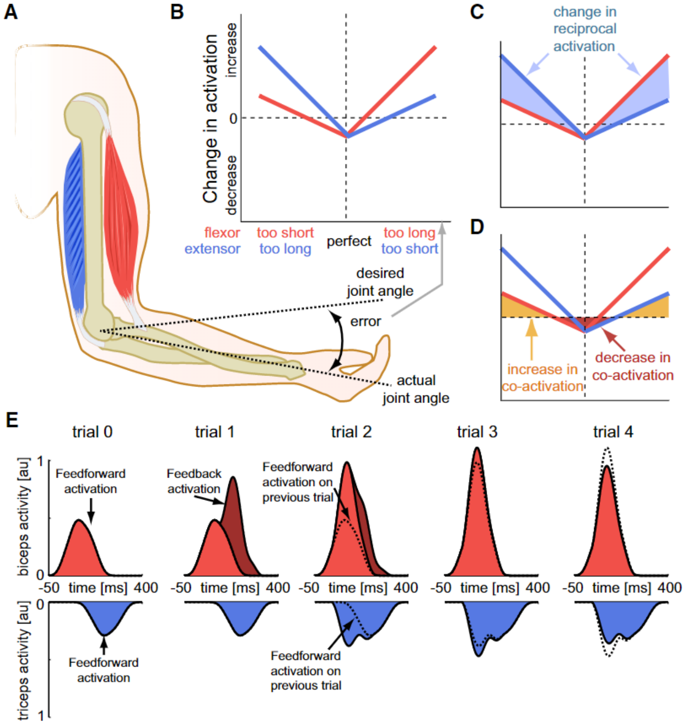
- $e_i^k(t)$: stretch/shortening in muscle i at time t
- Δu_i^k is shifted forward in time by $\zeta > 0$
- ζ: feedback delay
- α, β, γ: as shown in slide 13

Error-Based Learning Algorithm

Computational Mechanisms of Sensorimotor Control, David W. Franklin et al.,Neuron , 2008

Consider single joint motion

- A:
  - elbow joint is controlled by two antagonist muscles, biceps and triceps.
- B:
  - each error measure used by V-shaped update rule to determine change in muscle activation for next repetition of movement,
  - change is shifted forward in time to compensate for delays,
  - V-shaped learning rule for each muscle has different slope depending on whether muscle too long or too short (indicated by error)
- C:
  - different slopes for stretch/shortening of each muscle lead to appropriate change in reciprocal muscle activation
  - drives compensatory changes in the joint torques and endpoint forces



Adaptive Impedance Control for a Manipulator

Motivation and Control Formulation

- Ultimate goal of impedance control is the following task space closed-loop behavior (see lecture 4 *Modern Methods of Robot Control*)

$$\boldsymbol{F}_{\text{ext}} = \boldsymbol{M}\Delta\ddot{\boldsymbol{x}} + \boldsymbol{D}\Delta\dot{\boldsymbol{x}} + \boldsymbol{K}\Delta\boldsymbol{x}$$

- Control law:

$$\boldsymbol{\tau}_{\text{in}}(t) = \boldsymbol{J}^T(\boldsymbol{q})\big(-\boldsymbol{F}(t) - \boldsymbol{K}(t)\Delta\boldsymbol{x}(t) - \boldsymbol{D}(t)\Delta\dot{\boldsymbol{x}}(t)\big) + \boldsymbol{\tau}_r(t)$$

where

  - $\Delta\boldsymbol{x}(t) = \boldsymbol{x}(t) - \boldsymbol{x}_{\boldsymbol{d}}(t)$: position error
  - $\boldsymbol{F}(t)$: *feedforward* generalized force
  - $\boldsymbol{K}(t)$: *stiffness* matrix
  - $\boldsymbol{D}(t)$: *damping* matrix
  - $\boldsymbol{\tau}_r(t)$: rest of control law including gravity term
  - $\boldsymbol{J}(\boldsymbol{q})$: Jacobian matrix of the manipulator

Following the same idea as for human's limb, one can find an adaption law for the stiffness, damping and feedforward generalized force

Adaption Goal

Considering $\boldsymbol{\phi}(t)$ as the vector of the values which are to be adapted, and $\boldsymbol{\phi}^*(t)$ as the vector of desired values with which we can reach to the desired behaviour, our goal is to see that $\boldsymbol{\phi}(t)$ approaches $\boldsymbol{\phi}^*(t)$ while the error is being minimized

$$\boldsymbol{\phi}(t) \equiv [\mathrm{vec}(\boldsymbol{K}(t))^T, \mathrm{vec}(\boldsymbol{D}(t))^T, \boldsymbol{F}^T(t)]^T$$
$$\boldsymbol{\phi}^*(t) \equiv [\mathrm{vec}(\boldsymbol{K}_E(t))^T, \mathrm{vec}(\boldsymbol{D}_E(t))^T, \boldsymbol{F}_E^T(t)]^T$$

Considering $\tilde{\boldsymbol{\phi}}(t) \equiv \boldsymbol{\phi}(t) - \boldsymbol{\phi}^*(t)$, reaching to the goal would be equivalent to minimize the following two cost functions:

$$V_p(t) = \frac{1}{2}\boldsymbol{\varepsilon}^T(t)\boldsymbol{M}_C(\boldsymbol{q})\boldsymbol{\varepsilon}(t)$$
$$V_c(t) = \frac{1}{2}\tilde{\boldsymbol{\phi}}^T(t)\boldsymbol{Q}^{-1}(\boldsymbol{q})\tilde{\boldsymbol{\phi}}(t)$$

where

- $\boldsymbol{M}_C(\boldsymbol{q})$: mass matrix in task space
- $\boldsymbol{\varepsilon}(t) \equiv \Delta\dot{\boldsymbol{x}}(t) + \boldsymbol{\Lambda}\Delta\boldsymbol{x}(t)$: tracking error
- $\boldsymbol{Q}$: matrix made by the positive definite matrices $\boldsymbol{Q}_K, \boldsymbol{Q}_D$ and $\boldsymbol{Q}_\tau$, corresponding to the learning rate of $\boldsymbol{K}, \boldsymbol{D}$ and $\boldsymbol{\tau}$

Adaption Law

Minimization of $V(t) = V_p(t) + V_c(t)$ can be done by the adaption laws, similar to what we saw for human's limb

Adaption law for each time step can be defined as:

$$\boldsymbol{F}(t_{i+1}) = \boldsymbol{F}(t_i) + \boldsymbol{Q}_\tau(\boldsymbol{\varepsilon}(t_i) - \gamma(t_i)\boldsymbol{F}(t_i))$$
$$\boldsymbol{K}(t_{i+1}) = \boldsymbol{K}(t_i) + \boldsymbol{Q}_K(\boldsymbol{\varepsilon}(t_i)\Delta\boldsymbol{x}^T(t_i) - \gamma(t_i)\boldsymbol{K}(t_i))$$
$$\boldsymbol{D}(t_{i+1}) = \boldsymbol{D}(t_i) + \boldsymbol{Q}_D(\boldsymbol{\varepsilon}(t_i)\Delta\dot{\boldsymbol{x}}^T(t_i) - \gamma(t_i)\boldsymbol{D}(t_i))$$

where $\gamma(t_i)$ is the forgetting factor of learning which can be defined as follows

$$\gamma(t_i) = \frac{a}{1 + b\left|\left|\boldsymbol{\varepsilon}(t_i)\right|\right|^2}$$

with positive parameters a and b