

A motor control model

Thomas Beucher

June 12, 2015

Fiche d'évaluation du stage de fin d'études "Entreprise"
à compléter par le tuteur de stage
à retourner avec le rapport de stage **Obligatoirement**

Fiche d'évaluation du stage
Ing5 3^{ème} année Cycle Ingénieur
Promotion 2015

Etudiant : ...BEUCHER..... Majeure :
 Entreprise :UPMC - ISIR.....
 Tuteur de stage :Olivier SIGAUD.....
 ☐ du tuteur : ...01.44.27.88.53.... Email du tuteur : ...Olivier.Sigaud@upmc.fr

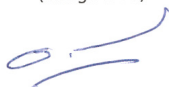
<i>Possibilité d'accorder des ½ points</i>		Note
Qualités Techniques		
• Connaissance des fondamentaux techniques		2/4
• Résultats obtenus		4/6
• Qualité du travail (méthodologie, réalisation, dossier ...)		7/10
Total Note Qualités Techniques		13/20
Qualités humaines		
• Autonomie et sens des responsabilités		4/7
• Esprit d'initiative et créativité		3/5
• Intégration / sociabilité au sein de l'équipe		4/5
• Comportement / Attitude		2/3
Total Note Qualités Humaines		13/20
TOTAL		26/40
Soit		13/20

Observations :

Le travail était essentiellement individuel. Le sujet de stage était difficile, dans un cadre recherche, il portait sur des notions dont Thomas BEUCHER n'avait pas la maîtrise au départ.

Compte-tenu de l'ambition du contexte, Thomas BEUCHER a été contraint de beaucoup progresser pour atteindre les objectifs fixés. Les objectifs ne sont pas encore complètement atteints au moment où je rédige cette évaluation, mais les 7 semaines restantes devraient suffire pour terminer le travail.

A ...Paris....., le ...8/6/2015
Signature du Tuteur de stage
 (obligatoire)



Cachet de l'entreprise d'accueil
 (obligatoire)

I.S.I.R
 Institut des Systèmes Intelligents & de Robotique
 Pyramide - Tour 55/56 - Boîte Courrier 173
 4, place Jussieu - 75252 PARIS Cedex 05
 Tél. 01 44 27 51 31/41 - Fax 01 44 27 51 45
www.isir.upmc.fr

Contents

1	Introduction	4
1.1	State of the art	5
1.2	Previous work	7
1.3	My job	7
2	Material and methods	8
2.1	Arm model	8
2.1.1	Arm parameters	8
2.1.2	Muscles parameters	9
2.1.3	Rigid-body dynamics	10
2.1.4	Muscular actuation	11
2.2	Mathematical formulation of the model	11
2.3	Learning method, RBFN controller	12
2.3.1	Theory	12
2.3.2	Implementation	14
2.4	Optimisation, CMAES controller	14
3	Study	16
3.1	Experimental set-up	16
3.2	Experiments	18
3.2.1	Model tests	18
3.2.2	RBFN controller	18
3.2.3	Optimize the RBFN controller	18
3.3	Results	19
3.3.1	Performance of the RBFN controller	19
3.3.2	Performance of the CMA-ES controllers	20
4	Appendix	27
4.1	Nomenclature of arm parameters	27
4.2	Experiment parameters	29
4.3	RBFN implementation	30
4.4	Experimental methods on human	34

Chapter 1

Introduction

This report is about my internship at ISIR covering a period of five months from February to July 2015.



The Institute for Intelligent Systems and Robotics (ISIR) is a multidisciplinary research laboratory that brings together researchers and academics from different disciplines of Engineering Sciences and Information and the Life Sciences. It is based in Paris France.

I choose this internship because I wanted to experience the world of research in view of prepare for the Phd that I hope to achieve. So I join the team of Olivier Sigaud to discover the field of computational neuroscience, especially the motor learning (what is the brain strategy to lead our movements).

1.1 State of the art

There has been a recent progress in motor control research on understanding how the time of a reaching movement is chosen. In particular, two recent models from Shadmehr et al. [17] and Rigoux&Guigon [15] proposed an optimization criterion that involves a trade-off between the muscular effort and the subjective value of getting the reward, hence a cost-benefit trade-off (CBT). On one hand, reaching a target faster requires a larger muscular effort. On the other hand, the subjective value of reaching a target decreases as the time needed to reach the target is increased. As a result, the net expected return consisting of the subjective value minus the muscular effort is optimal for a certain time, as illustrated in Fig. 1.1(A).

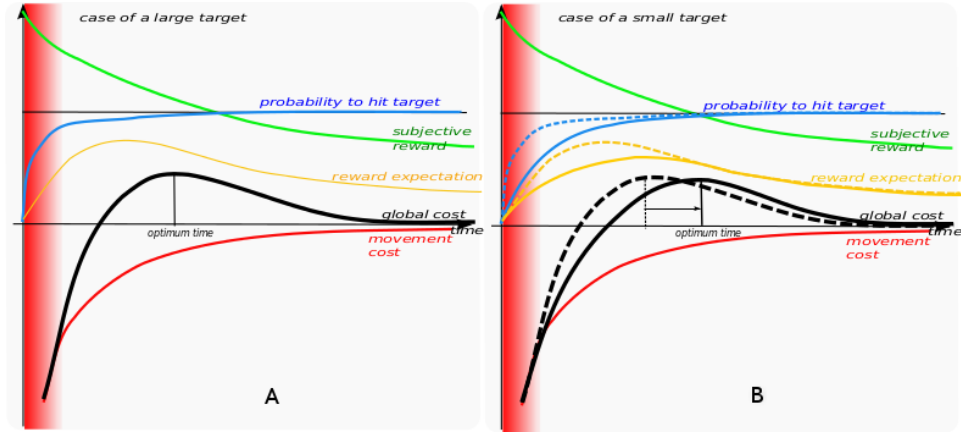


Figure 1.1: Influence of movement time on cost related quantities. Green: subjective utility of hitting the target; red: muscular energy cost; black: global cost versus reward trade-off. The red area denotes infeasible short times; blue: probability to hit the target; orange: reward expectation (subjective reward times probability). A: Sketch of the models in [17] and [15]. The subjective utility of hitting the reward decreases over time as one is less interested in gains that will occur in a distant future than at the present time. Hitting is less and less costly in terms of efforts as the movement is performed more slowly. The expected gain, resulting from the sum of the subjective reward and the (negative) cost reaches a maximum for a certain time. When the gain is negative (outside the useful interval), one should not move. B: Sketch of the presented model. In the case of a larger target, the hitting probability is higher for faster movements (solid lines) than for a smaller target (dashed lines). As a result, the maximum of the reward expectation is shifted towards longer time for smaller targets, and the optimum movement time is also longer for smaller targets.

However, these models do not account directly for basic facts about the

relation between movement difficulty and movement duration as captured more than fifty years ago by Fitts' law [6]. According to this law, the smaller a target, the slower the reaching movement. This is well explained by the so-called *speed-accuracy trade-off* (SAT) stating that, the faster a movement, the less accurate it is, hence the higher the probability to miss the target. So a subject reaching too fast may not get the subjective value associated to reaching and should slow down.

In contrast with the models of [17] and [15], the model of Dean [4] takes the SAT into account. The key difference with respect to [17] and [15] is that, instead of maximizing a reward, this model maximizes a *reward expectation*, i.e. the reward times the probability to get it.

However, the model proposed in [4] is an abstract model of movement time selection that looks for an optimal trade-off between an externally decayed reward and a SAT that relates the probability of missing to movement time. As such, it does not account for movement execution, neither for the choice of a motor trajectory and its impact on the cost of movement. The model does not explain Fitts' law, it rather incorporates its consequences into an abstract model of the SAT that is fitted to experimental data. The mathematical design of the model is based on several simplifying assumptions and it predicts optimal movement times that are systematically shorter than those observed with subjects. The authors of [4] discuss that this may result from the fact that the model does not take the cost of movement into account.

The goal of my work is to show that the models of Shadmehr [17] and Rigoux [15] as well as the model of Dean [4] can be unified into a model that solves the difficulties faced by these previous models.

This unification is simply implemented by including sensory and motor noise into the optimal control model proposed in [15], shifting from a deterministic account of the movement to a stochastic one, in line with the models of [7, 22, 20, 21, 11].

As a matter of fact, in the models of [17] and [15], the target is given as a single point and the movement is considered as always reaching it, irrespective of the size of the target. In order to fully account for Fitts' law, one must consider the intrinsic dispersion of reaching movements towards a target and the effect of sensory and muscular noise on this dispersion (e.g. [7], see [5] for a review), which is not the case of the models of [17] and [15].

As a result, the reward and muscular activation terms in the optimization criterion proposed in [15] are replaced by reward and cost expectation terms. Considering expectation is a way to account for the fact that, in case of a miss, one would not get the reward, so the global outcome of the movement would only consist of its incurred cost.

1.2 Previous work

I used a code, written by Kevin Monfray (a trainee at ISIR), which implements a model based on optimal control called NOPS but it is very time consuming. To reduce the computation time, Olivier Sigaud decided to try to learn the NOPS controller using a regression algorithm. So I generated trajectories using the NOPS controller then I used the regression algorithm called RBFN (Radial Basis Function Networks) to learn the new controller from these trajectories. The new controller could then be optimized with a stochastic optimization method called CMA-ES (Covariance Matrix Adaptation Evolution Strategy).

1.3 My job

First, I have to implement in python the model of arm and the regression algorithm to learn the new controller from the trajectories generated with the NOPS controller (implemented by Didier Marin in C++ and in Java by Kevin Monfray). Secondly, I must use the CMAES algorithm to optimize RBFN controller for different targets size. Finally, I must generate various relevant curves and compare them with the curves obtained by Slovenian colleagues, Jan babic and Luka Peternel, who experiment on humans around the same issue.

Chapter 2

Material and methods

2.1 Arm model

The plant is a two degrees-of-freedom (dofs) planar arm controlled by 6 muscles, illustrated in Fig. 2.1. There are several such models in the literature. The model described in [8] lies in the vertical plane so it takes the gravity force into account. Most other models are defined in the sagittal plane and ignore gravity effects. They all combine a simple two dofs planar rigid-body dynamics model with a muscular actuation model. The differences between models mostly lie in the latter component.

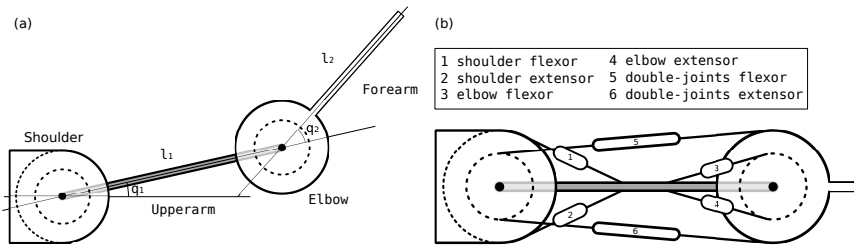


Figure 2.1: Arm model. (a) Schematic view of the arm mechanics. (b) Schematic view of the muscular actuation of the arm, where each number represents a muscle whose name is in the box.

Table 4.1 in Appendix 4.1 reminds the nomenclature of all the parameters and variables of the arm model.

2.1.1 Arm parameters

All parameters of the arm are defined in the file *setupArmParameters.py* and implemented in the class *ArmParameters*. This class defines the following functions:

readSetupFile : Reads the setup file.

massMatrix : Defines the inertia matrix parameters.

BMatrix : Defines the damping matrix \mathbf{B} , with $\mathbf{B} = \begin{bmatrix} .05 & .025 \\ .025 & .05 \end{bmatrix} \dot{q}$.

AMatrix : Defines the moment arm matrix \mathbf{A} .

$$\begin{aligned} \mathbf{A}^\top &= \begin{bmatrix} a_1 & a_2 & a_3 & a_4 & a_5 & a_6 \\ a_7 & a_8 & a_9 & a_{10} & a_{11} & a_{12} \end{bmatrix} \\ &= \begin{bmatrix} .04 & -.04 & 0 & 0 & .028 & -.035 \\ 0 & 0 & .025 & -.025 & .028 & -.035 \end{bmatrix} \end{aligned}$$

All the arm parameters values are summarized in Table 4.2 in Appendix 4.1.

2.1.2 Muscles parameters

All muscles parameters are defined in the file *setupMusclesParameters* and implemented in the class *MusclesParameters*. This class defines the following functions:

fmaxMatrix : Defines the matrix of the maximum force exerted by each muscle.

$$\begin{aligned} \mathbf{f}_{\max} &= \begin{pmatrix} f_{\max 1} & 0 & 0 & 0 & 0 & 0 \\ 0 & f_{\max 2} & 0 & 0 & 0 & 0 \\ 0 & 0 & f_{\max 3} & 0 & 0 & 0 \\ 0 & 0 & 0 & f_{\max 4} & 0 & 0 \\ 0 & 0 & 0 & 0 & f_{\max 5} & 0 \\ 0 & 0 & 0 & 0 & 0 & f_{\max 6} \end{pmatrix} \\ &= \begin{pmatrix} 700 & 0 & 0 & 0 & 0 & 0 \\ 0 & 382 & 0 & 0 & 0 & 0 \\ 0 & 0 & 572 & 0 & 0 & 0 \\ 0 & 0 & 0 & 445 & 0 & 0 \\ 0 & 0 & 0 & 0 & 159 & 0 \\ 0 & 0 & 0 & 0 & 0 & 318 \end{pmatrix} \end{aligned}$$

activationVectorInit : Initializes the muscular activation vector. (Creates the vector and initializes to zero)

activationVectorUse : Builds the muscular activation vector given its 6 components.

All the muscles parameters values are summarized in Table 3 in Appendix 4.1.

2.1.3 Rigid-body dynamics

The rigid-body dynamics equation of a mechanical system is:

$$\ddot{\mathbf{q}} = \mathbf{M}(\mathbf{q})^{-1}(\boldsymbol{\tau} - \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}}) - \mathbf{g}(\mathbf{q}) - \mathbf{B}\dot{\mathbf{q}}) \quad (2.1)$$

where \mathbf{q} is the current articular position, $\dot{\mathbf{q}}$ the current articular speed, $\ddot{\mathbf{q}}$ the current articular acceleration, \mathbf{M} the inertia matrix, \mathbf{C} the Coriolis force vector, $\boldsymbol{\tau}$ the segments torque, \mathbf{g} the gravity force vector and \mathbf{B} a damping term that contains all unmodelled effects. Here, \mathbf{g} is ignored since the arm is working in the sagittal plane. All angles are expressed in radians. We can compute the inertia matrix as: $\mathbf{M} = \begin{bmatrix} k_1 + 2k_2 \cos(q_2) & k_3 + k_2 \cos(q_2) \\ k_3 + k_2 \cos(q_2) & k_3 \end{bmatrix}$, with $k_1 = d_1 + d_2 + m_2 l_1^2$, $k_2 = m_2 l_1 s_2$, $k_3 = d_2$ where d_i , m_i , l_i and s_i are parameters of the arm previously defined in Section 2.1.1.

The Coriolis force vector is given by

$$\mathbf{C} = \begin{bmatrix} -\dot{q}_2(2\dot{q}_1 + \dot{q}_2)k_2 \sin(q_2) \\ \dot{q}_1^2 k_2 \sin(q_2) \end{bmatrix}.$$

The computation of the torque τ exerted on the system given an input muscular actuation \mathbf{u} is explained in the section 2.1.4.

Equation 2.1 is implemented in the class *ArmDynamics* line 57:

```
57     ddotq = np.dot(Minv, (Gamma - C - np.dot(armP.B,
        dotq)))
```

where *np* refers to the numpy library in python. We also find in this class all elements of Equation 2.1:

```
44     #Inertia matrix
45     M = np.array([[armP.k1+2*armP.k2*math.cos(q[1,0]),
        armP.k3+armP.k2*math.cos(q[1,0])],[armP.k3+armP
        .k2*math.cos(q[1,0]),armP.k3]])
46     #Coriolis force vector
47     C = np.array([[-dotq[1,0]*(2*dotq[0,0]+dotq[1,0])*
        armP.k2*math.sin(q[1,0])],[ (dotq[0,0]**2)*armP.
        k2*math.sin(q[1,0])]])
48     #inversion of M
49     Minv = np.linalg.inv(M)
50     #torque term
51     Q = np.diag([q[0,0], q[0,0], q[1,0], q[1,0], q
        [0,0], q[0,0]])
52     #the commented version uses a non null stiffness
        for the muscles
53     #Gamma = np.dot((np.dot(armP.At, musclesP.fmax)-np
        .dot(musclesP.Kraid, Q)), U)
```

54 `Gamma = np.dot((np.dot(armP.At, musclesP.fmax)-np.
 dot(musclesP.Knulle, Q)), U)`

2.1.4 Muscular actuation

Our muscular actuation model is taken from [9] (pp. 356-357) through [12]. It is a simplified version of the one described in [11] in the sense that it uses a constant moment arm matrix \mathbf{A} whereas [11] is computing this matrix as a function of the state of the arm.

Finally, given an action \mathbf{u} corresponding to a raw muscular activation as output of the controller, the muscular activation is augmented with Gaussian noise using $\tilde{\mathbf{u}} = \log(\exp(\kappa \times \mathbf{u}_t \times (1 + \mathcal{N}(0, \mathbf{I}\sigma_u^2))) + 1)/\kappa$, where \times refers to the element-wise multiplication, \mathbf{I} is a 6×6 identity matrix. and $\kappa = 25$ is the Heaviside filter parameter, and the input torque is computed as $\boldsymbol{\tau} = \mathbf{A}^\top(\mathbf{f}_{\max} \times \tilde{\mathbf{u}})$.

2.2 Mathematical formulation of the model

The cost function $J(\vec{u})$ proposed for a control \vec{u} in the model of [15] is

$$J(\vec{u}) = \int_0^\infty e^{-t/\gamma} [\rho R(\vec{s}_t) - \nu L(\vec{u}_t)] dt \quad (2.2)$$

where $R(\vec{s}_t)$ is the immediate reward function that equals 1 at the goal point (also called rewarded state) and is null everywhere else. The function $L(\vec{u}_t)$ is the movement cost. The authors of [15] take $L(\vec{u}_t) = \|\vec{u}_t\|^2$, as in many motor control models. The continuous-time discount factor γ accounts for the “greediness” of the controller, i.e. the smaller γ , the more the agent is focused on short term rewards. Finally, ρ is the weight of the reward term and ν the weight of the effort term. In all experiments presented here, based, on the previous work from [15], we took $\gamma = 0.998$, $\rho = 300$ and $\nu = 1$.

A near optimal deterministic policy to solve this problem is obtained through a computationally expensive variation calculus method (see [14] for details). Given that the policy does not take the presence of noise in the model of the plant into account, the actions must be computed again at each time step depending on the new state reached by the plant which further contributes to the cost of the method. The controller resulting from this model is called the NOPS (for Near-Optimal Planning System).

Now let us consider the integration of accuracy constraints. Instead of a deterministic controller, the new model is based on a stochastic controller where the rewarded state is reached or not. As a result, the outcome of a large set of movements performed with noise is computed as the value of the reward multiplied by the probability to obtain it over the different

movements. Mathematically, the value multiplied by the probability is called the expectation.

Taking the probability to reach the target into account as described above, the new optimization criterion is written

$$J(\vec{u}) = \int_0^\infty e^{-t/\gamma} \mathbb{E}[\rho R(\vec{s}_t) - \nu L(\vec{u}_t)] dt \quad (2.3)$$

where $\mathbb{E}[\cdot]$ stands for the expectation of the cumulated reward, and $R(\vec{s}_t)$ equals 1 if the end effector hits the target.

2.3 Learning method, RBFN controller

2.3.1 Theory

Regression is the process of learning relationships between inputs and continuous outputs from example data, which enables predictions for novel inputs. The history of regression is closely related to the history of artificial neural networks since the seminal work of Rosenblatt (1958). There are many regression algorithms, in our case we use the *Regression with Radial Basis Function Networks* (RBFNs).

We want to approximate the function $f(s) = U$ where s is the state and U the vector of muscular activations. We use a parametric function $\hat{f}_\theta : \mathbb{R} \rightarrow \mathbb{R}$ where $\theta \in M_{k,n}(\mathbb{R})$ is a real matrix of parameters. We make the assumption that f is non-linear and can be approximated using:

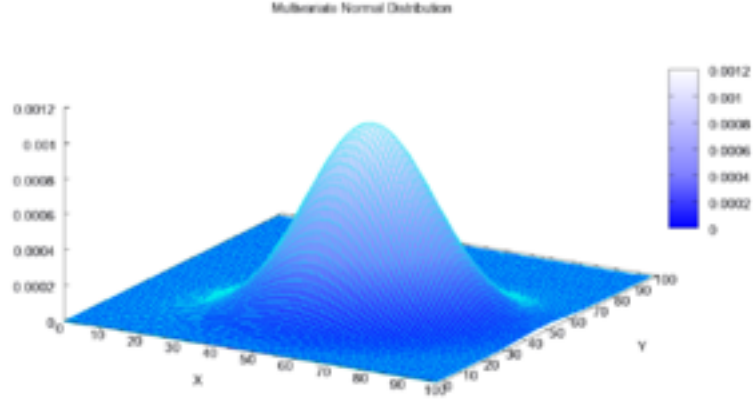
$$\hat{f}(s) = \Phi(s)^T \theta \quad (2.4)$$

where $\phi(s) = [\phi_1(s), \phi_2(s), \dots, \phi_k(s)]^T$ is a vector function commonly referred as features.

We use Multivariate Gaussian basis functions throughout the input space. We are in the non-degenerate case i.e. when the symmetric covariance matrix Σ is positive definite. The multivariate normal distribution in our case could be written as:

$$f_x(s_1, \dots, s_k) = \frac{1}{\sqrt{(2\pi)^k |\Sigma|}} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu)\right) \quad (2.5)$$

where x is a real k -dimensional column vector and $|\Sigma|$ is the determinant of the covariance matrix Σ . The descriptive statistics $(x - \mu)^T \Sigma^{-1}(x - \mu)$ in the non-degenerate multivariate normal distribution equation is known as the square of the Mahalanobis distance, which represents the distance of the test point x from the mean μ .



The goal is to determine the parameters θ which permit \hat{f} to best approximate f . Let ∇_{θ} denote the gradient according to θ , the gradient of the error function is:

$$\begin{aligned}
 \nabla_{\theta} \hat{\epsilon}(\theta) &= \nabla_{\theta} \mathbb{E} \left[\frac{1}{2} (y - \hat{f}(s))^2 \right] \approx \nabla_{\theta} \frac{1}{N_s} \sum_{i=1}^{N_s} \left[\frac{1}{2} (y - \hat{f}(s))^2 \right] \\
 &= \frac{1}{N_s} \sum_{i=1}^{N_s} \left[(\nabla_{\theta} (y - \hat{f}(s))) (y - \hat{f}(s)) \right] \\
 &= \frac{1}{N_s} \sum_{i=1}^{N_s} \left[(\nabla_{\theta} \hat{f}(s)) (\hat{f}(s) - y) \right]. \quad (2.6)
 \end{aligned}$$

For our linear function approximator with Gaussian basis functions, we have

$$\nabla_{\theta} \hat{f}(s) = \phi(s). \quad (2.7)$$

This gradient cancels out for any local optimum of the error function, therefore we can find one of these by formulating θ as follows:

$$\begin{aligned}
 \nabla_{\theta} \hat{\epsilon}(\theta) &= 0 \\
 \Leftrightarrow \quad \frac{1}{N_s} \sum_{i=1}^{N_s} \phi(s_i) (\phi(s_i)^T \theta - y_i) &\approx 0 \\
 \Leftrightarrow \quad \sum_{i=1}^{N_s} (\phi(s_i) \phi(s_i)^T) \theta &\approx \sum_{i=1}^{N_s} (\phi(s_i) y_i) \\
 \Leftrightarrow \quad A \theta &\approx b \\
 \Leftrightarrow \quad \theta &\approx A^{\#} b
 \end{aligned}$$

with,

$$A = \sum_{i=1}^{Ns} (\phi(s_i)\phi(s_i)^T) \quad (2.8)$$

$$b = \sum_{i=1}^{Ns} (\phi(s_i)y_i) \quad (2.9)$$

where $A^\#$ is the pseudo-inverse of matrix A.

2.3.2 Implementation

We have a set of N_s samples (the data of trajectories generated by Monfray's controller given the state and the muscles activations at each time for each trajectories). We feed the algorithm with this data through the function *setTrainingData* which check the input dimension, the output dimension and if there is the same number of samples for input and output.

Then we set the centers and widths for the gaussian used by calling the function *setCentersAndWidths*.

Finally, we run the function *train_rbf* to create the controller. We find is this function:

- the computation of A line 46:

```
46         At = np.dot(fop, fop.T)
```

where $fop = \phi(s)$.

- the computation of b line 52:

```
52         bt = np.dot(fop, self.outputData.T)
```

- the computation of θ line 80:

```
80         self.theta = np.dot(np.linalg.pinv(A), b)
```

where $\text{np.linalg.pinv}(A) = A^\#$.

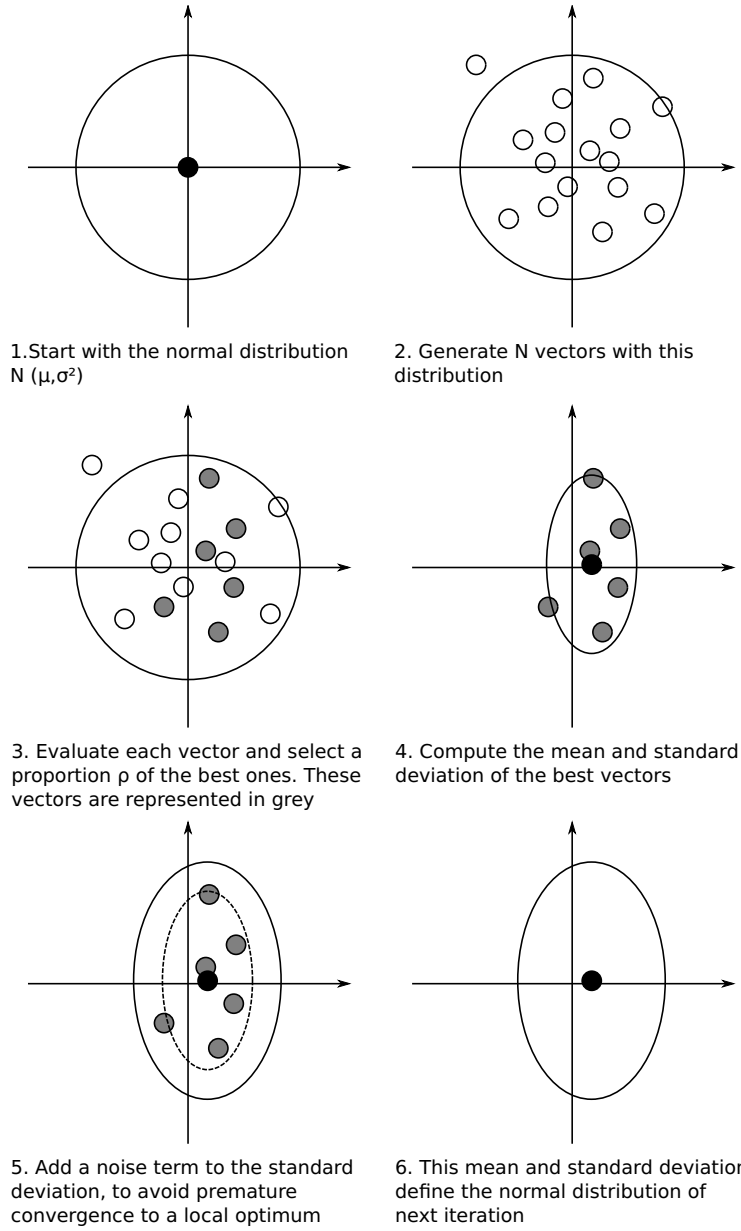
The code of the *fa_rbf* class is in appendix 4.3.

2.4 Optimisation, CMAES controller

After obtaining the RBFN controller, we want to optimize controllers for different sizes of target.

We use in this part the CMA-ES (Covariance Matrix Adaptation Evolution Strategy) algorithm implemented in the python module *cma*.

We can roughly explain how this algorithm work by the following scheme:



To launch the CMA-ES algorithm, we use the function *LaunchTrajectoriesCMAES* in *LaunchTrajectories.py*. This function run x times (here $x = 5$) each trajectories, evaluate them and return the average cost (here the cost is equal to the cost in term of muscle activations plus the reward if the target is reached). The algorithm feeds the function with the matrix θ and plays on each weight in order to maximize the obtained cost (i.e. to find the best compromise between the speed to reach the target and the accuracy).

Chapter 3

Study

3.1 Experimental set-up

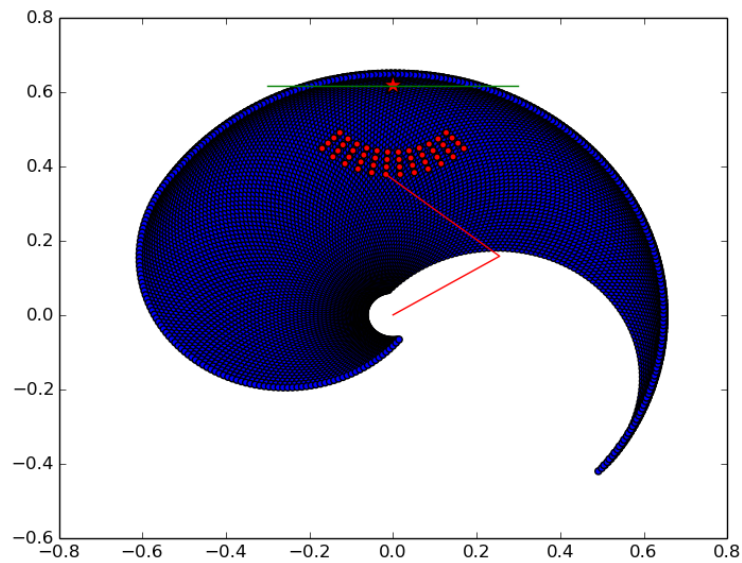


Figure 3.1: The arm workspace. The reachable space is delimited by a spiral-shaped envelope. The two segments of the arm are represented by two red lines. Initial movement positions are represented with red dots organized into four sets of different distances to the target. The screen is represented as a green line positioned at $y = 0.6175$. The origin of the arm is at $x = 0.0$, $y = 0.0$.

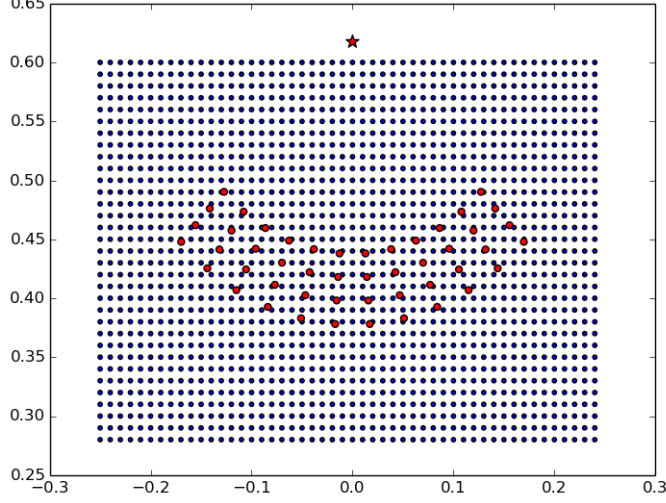


Figure 3.2: The blue dots represents the initial points of trajectories generated with NOPS. The red dot represents the initial positions used for our experiment. The red star represent the target.

The state-space consists of the current articular position \mathbf{q} of the arm and its current articular speed $\dot{\mathbf{q}}$. The state $\mathbf{s} = (\dot{\mathbf{q}}, \mathbf{q})$ has a total of 4 dimensions. The initial state is defined by null speed and a variable initial position. The positions are bounded to represent the reachable space of a standard human arm, with $q_1 \in [-0.6, 2.6]$ and $q_2 \in [-0.2, 3.0]$, as shown in Figure 3.1. The action-space consists of an activation signal for each muscle, which also makes a total of 6 dimensions.

When using CMAES, reaching the goal point is replaced by hitting the target on the screen. The target is defined as an interval of varying length around $(x = 0, y = 0.6175)$. The movement is stopped once the line $y = 0.6175$ has been crossed, and the intersect between the trajectory and this line is computed to determine whether the target was hit. The reward for immediately hitting the target without taking incurred costs into account is set to 300.

In order to train RBFN, we generate trajectories with the NOPS from the initial positions shown by the blue dots in Figure 3.2 to the goal point.

For our experiment, the starting points have been organized into four groups of different distances with respect to the goal point ($d = 18\text{cm}, 20\text{cm}, 22\text{cm}$ and 24cm respectively). There are 12 initial positions per distance, thus a total of 48 initial positions.

We measure the dispersion over 100 movements towards this target, as well as the average movement time and average movement cost.

We run 60 iterations of CMAES with a population of 100 individuals and $\sigma = 1.10^{-6}$ on the RBFN controller for each target.

For all the obtained controllers, we measure again the dispersion over 100 movements, the average movement time and average movement cost.

Finally, for all these targets, we record the velocity profile.

3.2 Experiments

All algorithms are implemented in Python. The parameters used for the experiments are given in the table 4.4 in Appendix.

3.2.1 Model tests

We now verify that our model behaves as expected. So we perform two experiments.

1. First, we create a set Γ of initial positions. Γ contains 48 initial positions organized into four sets of different distances to the target. Using the trajectories generated by the NOPS controller, we create the RBFN controller and evaluate its performance on each trajectories of Γ .
2. Second, for each target sizes, we perform x (here $x = 50$) times the reaching task for each initial positions of Γ . We measure different parameters during the generation of the trajectories:
 - coordinate of the elbow and the end-effector
 - articular speed of the shoulder and the elbow
 - the muscle activations
 - the cost of the trajectories (with and without the reward term)
 - the time taken to reach the goal

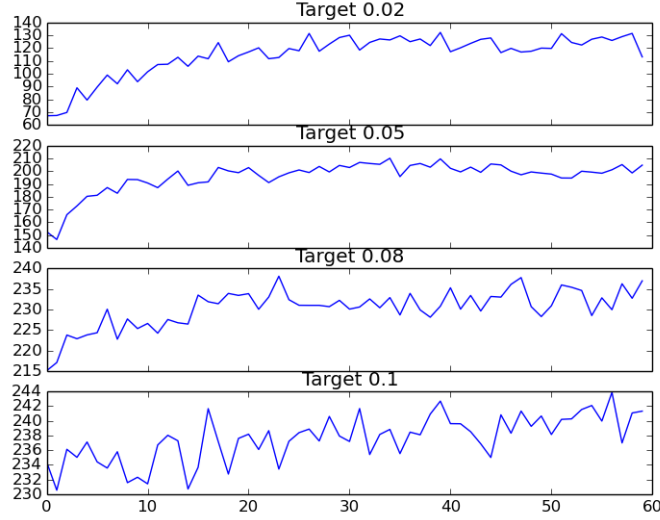
3.2.2 RBFN controller

First of all, we generate 1650 trajectories with the NOPS controller in order to train our new controller. (See the figure 3.2) Then, we evaluate the new controller obtained by looking at the cost map compared to the cost map of the NOPS controller.

3.2.3 Optimize the RBFN controller

Starting back from the best RBFN controller (best set of θ), we optimize, using the CMA-ES algorithm, for four different target's size. We specify the number of iterations required and the population size for CMA-ES. We

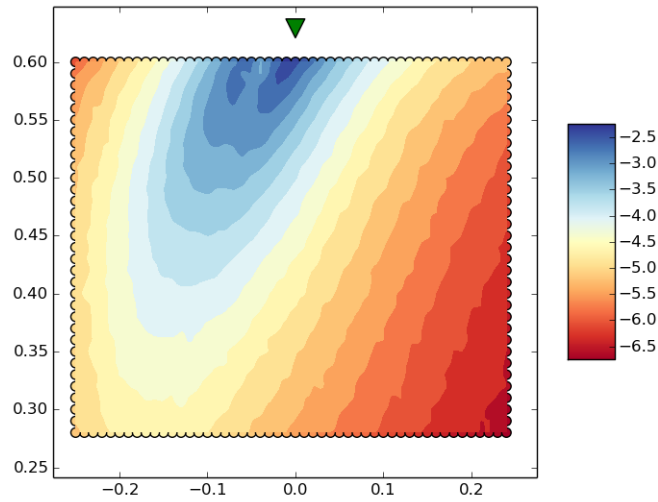
follow the evolution of the performance during the optimization. We now have four new controller optimized for each target's size and we can evaluate them to see if the results are consistent with the model.



3.3 Results

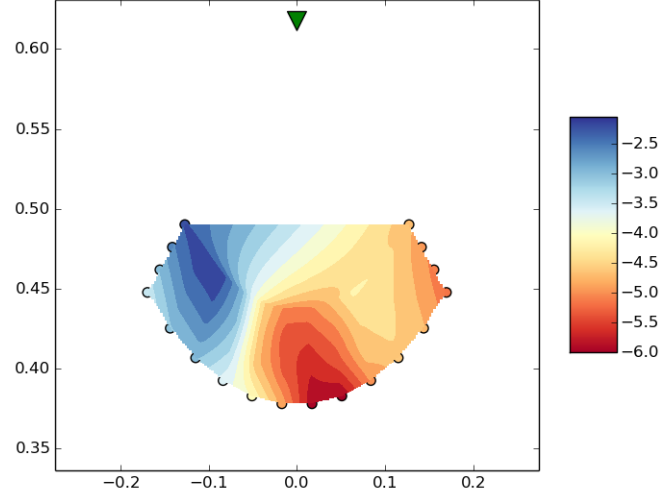
3.3.1 Performance of the RBFN controller

First, we evaluate the performance of the NOPS controller.



This figure show the cost (norm of muscle activations vector) of trajectories from different initial points. An area where the cost is low is observed. Moving away from this area, the cost trajectories gradually increase.

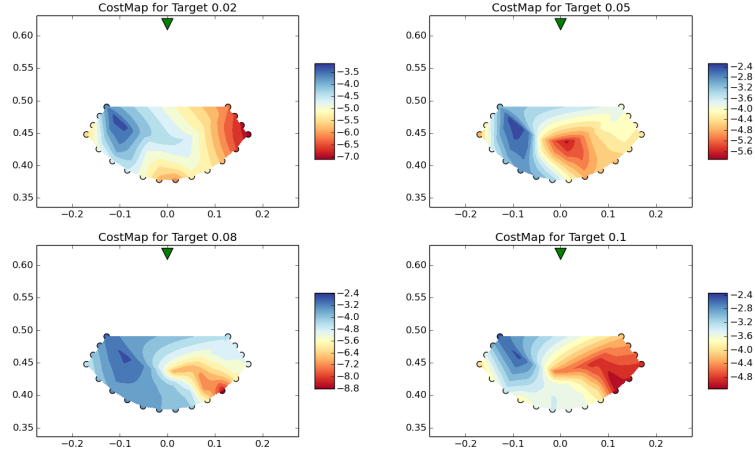
We evaluate the same things for the RBFN controller.



As expected, we can observe that starting from the left side of the goal point results in a lower cost than starting from the right side. This is explained by the fact that the optimal muscular strategy for performing these movements differs depending on the direction and the initial position.

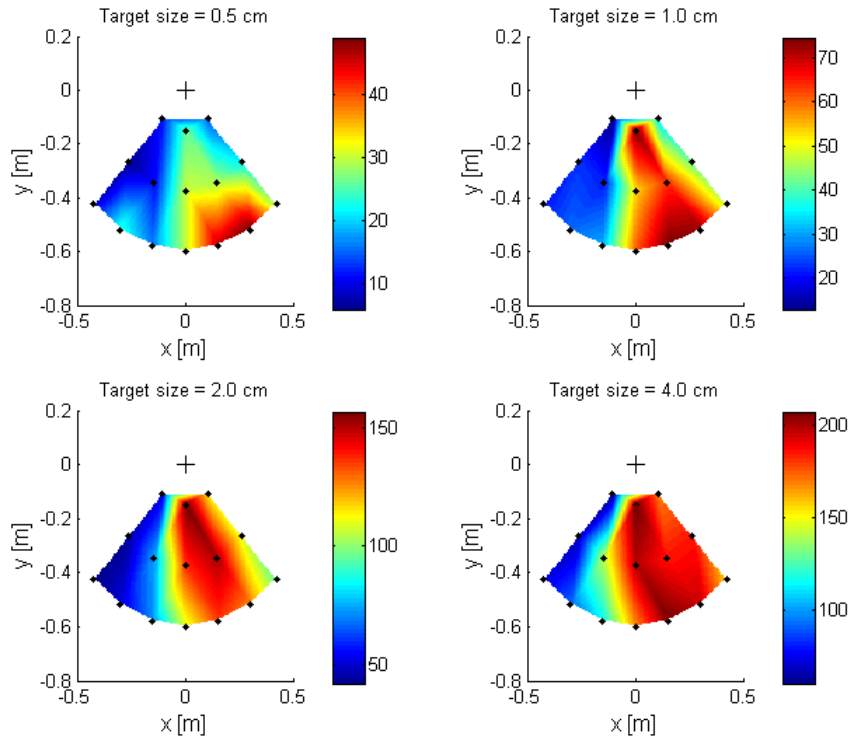
3.3.2 Performance of the CMA-ES controllers

First, we can evaluate the results of the optimization by looking at the cost map and if the strategy is still the same. We see that starting point from the left side of the goal point results in a lower cost than starting from the right side. So we plot the cost map obtained after evaluating the different controllers created by the optimization for four different targets:



We barely see the previous strategy but the area where the trajectories are the most costly is not as we expected in the right side. Maybe this is due to the strategy for muscle activation in this area because all the muscles have to be activated to ensure a good trajectory.

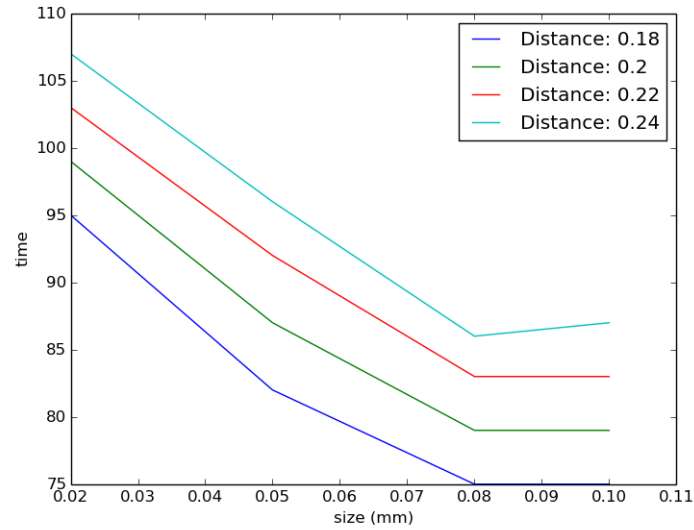
If we look at the cost map obtained by our Slovenian colleagues:



This map represent the cost calculated only from kinetic energy of the limb (no external force). This shows only energy related to the limb movement.

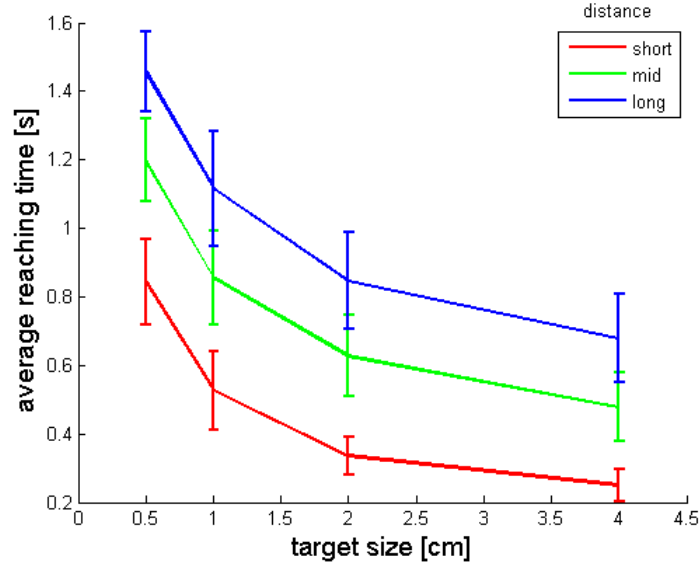
We find approximately the same pattern with a high cost area. The targets size are not the same in human experiment because we don't have the same accuracy than a computer and our model need smaller target sizes because otherwise the model always hits the target without any constraint on accuracy, so the laws do not emerge.

According to the theory there must be a trade-off between the speed to reach the target and the accuracy. We expect that to ensure to reach the target, when the size is reduced the speed must decrease at the end in order to be more accurate. In order to check this point, we plot the time to reach the target over the size of the target for different distances between the initial points and the target:



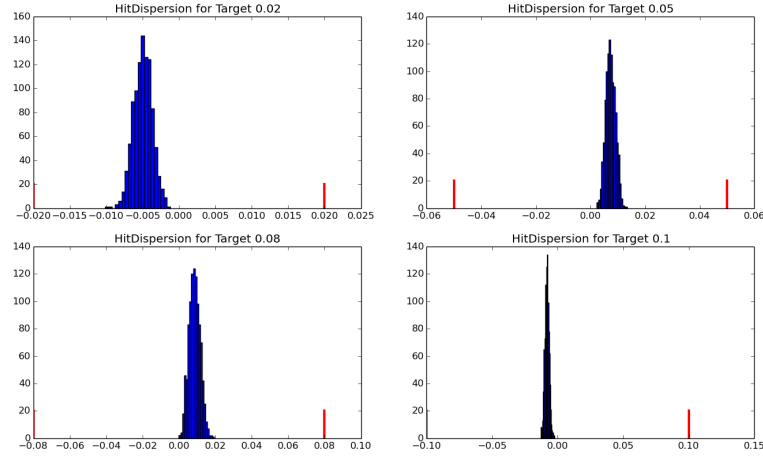
We can see that the time to reach the target increases with the distance to the target. We can also observe that the smaller the target, the more time taken to reach the target.

Our Slovenian colleagues obtained this figure:



We find the same pattern, the greater the distance to the target, the longer time to reach it. The time taken to reach the target decreases with the targets size.

We also expect that when the size of the target increases, the accuracy decreases over the speed to reach it. To analyse this part we can plot the hit dispersion:



This figure show the dispersion of the hit points. We see that the dispersion increases as much as the size of the target increases but not as much as we can expect. This can be explained by the fact that the accuracy is good enough at maximum speed so the arm does not miss the target.

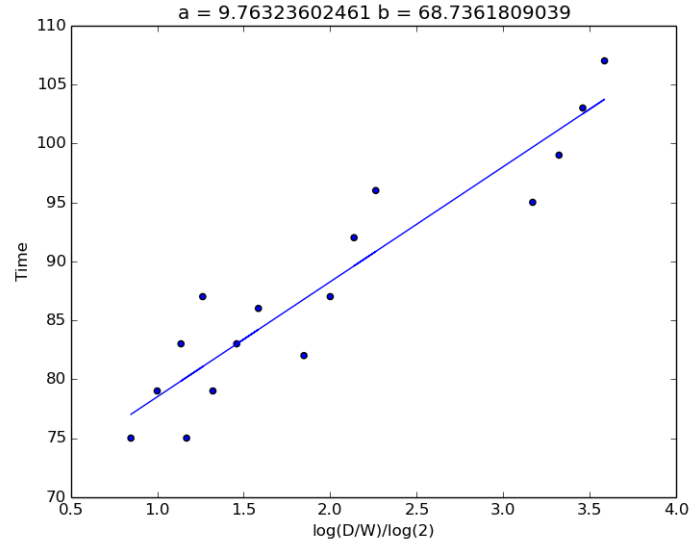
Reproduction of Fitt's law

Fitts' law states that movement time (MT) is linear in its difficulty index (DI), this index being bigger for longer movements and smaller targets. Fitts' law is written:

$$MT = a + b \cdot \underbrace{\log_2 \left(\frac{D}{W} \right)}_{DI} \quad (3.1)$$

where D is the distance of the movement, W is the width of the target and a and b are linear coefficients. This law was initially studied for one dimensional movements, and then extended for many other contexts [19, 2, 10, 13, 18, 1].

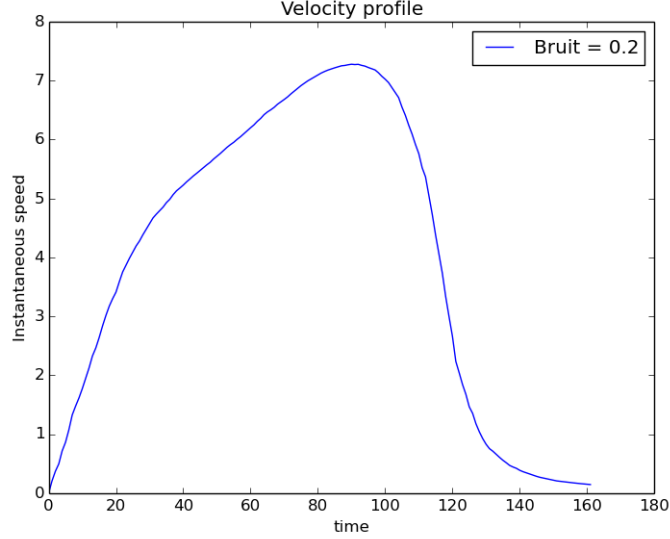
We compute DI values for different distances D and target widths W . the figure below shows the resulting movement time MT over DI .



One can see that we get a clear linear relationships, thus the data is consistent with Fitt's law.

The obtained values of a and b cannot be compared to empirical data from the human motor control literature given the wide variability of these values accross subjects [3, 16].

Velocity profile



This velocity profile is obtained with the smallest target, we can see that at the end the speed decreases a lot. It can be explained by the fact that to reach a small target, the controller must decrease the speed in order to increase the accuracy.

The final dispersion in reaching trajectories is generated by motor noise. Following the minimum intervention principle from [22], motor noise being proportional to muscular activation, the only way to decrease motor noise is to decrease muscular activation.

Thus, in order to hit a small trajectory, muscular activations should be small by the end of the movement, which can result in first instance in less co-contraction and then in less velocity. Furthermore, a slower movement provides a better opportunity for state estimation to compensate for delayed feedback about the position of the end effector. Taken together, those two phenomena contribute to the fact that an optimal controller should generate less velocity by the end of the movement for a smaller target. So one way to make sure to hit a small target would be to perform a slow reaching movement.

However, as explained above, a slower movement results in a discounted reward, thus the movement should nevertheless be as fast as possible.

As a consequence, the best option for optimizing reaching accuracy under temporal constraints consists in being very fast in the beginning of the movement and much slower in the end. Thus the velocity profile should be asymmetric. The main drive for this asymmetry being motor noise, the more motor noise, the more asymmetric the movement should be. Incidentally, one can observe on the ascending parts of the profiles that we do not

get strictly the shape of a bell curve.

At this point of my internship, I have not yet implemented the delayed feedback about the position of the end effector so I can not show this part of the experimentation.

Chapter 4

Appendix

4.1 Nomenclature of arm parameters

Table 4.1: Parameters of the arm model.

m_i	mass of segment i (kg)
l_i	length of segment i (m)
s_i	inertia of segment i ($kg.m^2$)
d_i	distance from the center of segment i to its center of mass (m)
κ	Heaviside filter parameter
\mathbf{A}	moment arm matrix ($\in \mathbb{R}^{6 \times 2}$)
\mathbf{f}_{\max}	maximum muscular tension ($\in \mathbb{R}^6$)
\mathbf{M}	inertia matrix ($\in \mathbb{R}^{2 \times 2}$)
\mathbf{C}	Coriolis force ($N.m \in \mathbb{R}^2$)
τ	segments torque ($N.m \in \mathbb{R}^2$)
\mathbf{B}	damping term ($N.m \in \mathbb{R}^2$)
\mathbf{u}	raw muscular activation (action) ($\in [0, 1]^6$)
σ_u^2	multiplicative muscular noise ($\in [0, 1]^6$)
\tilde{u}	filtered noisy muscular activation ($\in [0, 1]^6$)
\mathbf{q}^*	target articular position ($rad \in [0, 2\pi]^2$)
\mathbf{q}	current articular position ($rad \in [0, 2\pi]^2$)
$\dot{\mathbf{q}}$	current articular speed ($rad.s^{-1}$)
$\ddot{\mathbf{q}}$	current articular acceleration ($rad.s^{-2}$)

Table 4.2: Parameters of the arm.

l_1	arm length (m)	0.3
l_2	forearm length (m)	0.35
m_1	arm mass (kg)	1.4
m_2	forearm mass (kg)	1.1
s_1	arm inertia ($kg.m^2$)	0.11
s_2	forearm inertia ($kg.m^2$)	0.16
d_1	distance from the center of segment 1 to its center of mass (m)	0.025
d_2	distance from the center of segment 2 to its center of mass (m)	0.045
k_6	damping term	0.05
k_7	damping term	0.025
k_8	damping term	0.025
k_9	damping term	0.05
a_1	moment arm matrix	0.04
a_2	moment arm matrix	-0.04
a_3	moment arm matrix	0.0
a_4	moment arm matrix	0.0
a_5	moment arm matrix	0.028
a_6	moment arm matrix	-0.035
a_7	moment arm matrix	0.0
a_8	moment arm matrix	0.0
a_9	moment arm matrix	0.025
a_{10}	moment arm matrix	-0.025
a_{11}	moment arm matrix	0.028
a_{12}	moment arm matrix	-0.035

Table 4.3: Parameters of the muscles.

$f_{\max 1}$	Maximum force exerted by the shoulder flexor	700
$f_{\max 2}$	Maximum force exerted by the shoulder extensor	382
$f_{\max 3}$	Maximum force exerted by the elbow flexor	572
$f_{\max 4}$	Maximum force exerted by the elbow extensor	445
$f_{\max 5}$	Maximum force exerted by the double-joints flexor	159
$f_{\max 6}$	Maximum force exerted by the double-joints extensor	318

4.2 Experiment parameters

Table 4.4: Parameters of the experiment.

Number of features	3
Value of the motor noise	0.2
γ parameter for the cost function	0.998
ρ parameter for the cost function	300
v parameter for the cost function	1
σ parameter for CMA-ES	1.10^{-6}
Number of iteration for CMA-ES	60
Size of the population for CMA-ES	150
Targets size for the experiment (mm)	0.02/0.05/0.08/0.1
Time step (s)	0.002

4.3 RBFN implementation

```
15 class fa_rbfnn():
16
17     def __init__(self, nbFeature):
18         '''
19         Initializes class parameters
20
21         Input:      -nbFeature: int, number of feature
                       in order to perform the regression
22
23         '''
24         self.nbFeat = nbFeature
25         self.title = "rbfn"
26
27     def setTrainingData(self, inputData, outputData):
28         '''
29         Verifies the validity of the given input and
30         output data
31         Data should be organize by columns
32
33         Input:      -inputdata, numpy N-D array
34                     -outputData, numpy N-D array
35
36         '''
37         self.inputData = inputData
38         self.outputData = outputData
39         #Getting input and output dimensions and
40         number of samples
41         self.inputDimension, numberOfInputSamples = np
42         .shape(inputData)
43         self.outputDimension, numberOfOutputSamples =
44         np.shape(outputData)
45         #check if there are the same number of samples
46         for input and output data
47         assert(numberOfInputSamples ==
48                numberOfOutputSamples), "Number of samples
49                not equal for output and input"
50         self.numberOfSamples = numberOfInputSamples
51         self.theta = np.zeros((self.nbFeat, self.
52                                outputDimension))
53
54     def computeA(self, A, fop):
55         At = np.dot(fop, fop.T)
```

```

47         for i in range(At.shape[0]):
48             for j in range(At.shape[1]):
49                 A[i,j] = At[i,j]
50
51     def computeb(self, b, fop):
52         bt = np.dot(fop, self.outputData.T)
53         for i in range(bt.shape[0]):
54             for j in range(bt.shape[1]):
55                 b[i,j] = bt[i,j]
56
57     def train_rbfnn(self):
58         '''
59         Defines the training function to find solution
60         of the approximation
61
62         '''
63         fop = self.computeFeatureWeight(self.inputData
64                                         .T)
65         n = self.nbFeat**self.inputDimension
66         #creation of the shared objects between
67         process
68         AshareObj = Array(ct.c_double, n*n)
69         bshareObj = Array(ct.c_double, n*self.
70                           outputDimension)
71         #link shared object to a numpy object
72         AnumpyShare = np.frombuffer(AshareObj.get_obj
73                                     ())
74         bnumpyShare = np.frombuffer(bshareObj.get_obj
75                                     ())
76         #Reshaping of the numpy object to obtain an
77         array with the dimension desired
78         A = AnumpyShare.reshape((n, n))
79         b = bnumpyShare.reshape((n, self.
80                                 outputDimension))
81         #creation of the different processes
82         p1 = Process(target=self.computeA, args=(A,
83                                                  fop))
84         p2 = Process(target=self.computeb, args=(b,
85                                                  fop))
86         p1.start()
87         p2.start()
88         p1.join()
89         p2.join()
90         self.theta = np.dot(np.linalg.pinv(A), b)

```



```

81
82     def setCentersAndWidths(self):
83         '''
84         Sets the centers and widths of Gaussian
            features.
85         Uses linspace to evenly distribute the
            features.
86
87         '''
88         #get max and min of the input data
89         minInputData = np.min(self.inputData, axis =
            1)
90         maxInputData = np.max(self.inputData, axis =
            1)
91         rangeForEachDim = maxInputData - minInputData
92         #set the sigmas
93         widthConstant = rangeForEachDim / self.nbFeat
94         #create the diagonal matrix of sigmas to
            compute the gaussian
95         self.widths = np.diag(widthConstant)
96         #coef for gaussian
97         self.norma = 1/np.sqrt(((2*np.pi)**self.
            inputDimension)*np.linalg.det(self.widths))
98         linspaceForEachDim = []
99         #set the number of gaussian used and allocate
            them in each dimensions
100        for i in range(self.inputDimension):
101            linspaceForEachDim.append(np.linspace(
                minInputData[i], maxInputData[i], self.
                nbFeat))
102        #get matrix with all the possible combinations
            to find each centers
103        self.centersInEachDimensions = cartesian(
            linspaceForEachDim)
104
105    def computeFeatureWeight(self, inputData):
106        '''
107        Computes Gaussian parameters
108
109        Input:      -inputData: numpy N-D array
110
111        Output:     -phi: numpy N-D array
112        '''
113        #if only one sample

```

```

114         if inputData.shape[1] == 1:
115             x_u = inputData - self.
                    centersInEachDimensions.T
116             x_u_s = np.dot(x_u.T, np.linalg.pinv(self.
                    widths))
117             x = x_u_s * (x_u.T)
118             xf = np.sum(x, axis = 1)
119             phi = self.norma*np.exp(-0.5*xf)
120         #if more than one sample
121         else:
122             for i in range(inputData.shape[0]):
123                 x_u = np.array([inputData[i]]).T -
                    self.centersInEachDimensions.T
124                 x_u_s = np.dot(x_u.T, np.linalg.pinv(
                    self.widths))
125                 x = x_u_s * (x_u.T)
126                 xf = np.sum(x, axis = 1)
127                 xfe = self.norma*np.exp(-0.5*xf)
128                 if i == 0:
129                     phi = np.array([xfe]).T
130                 else:
131                     phi = np.hstack((phi, np.array([
                        xfe]).T))
132         return phi
133
134
135     def computesOutput(self, inputData, theta):
136         '''
137         Returns the output depending on the given
            input and theta
138
139         Input:         -inputData: numpy N-D array
140                        -theta: numpy N-D array
141
142         Output:        -fa_out: numpy N-D array, output
            approximated
143         '''
144         if inputData.shape[1] == 1:
145             phi = self.computeFeatureWeight(inputData)
146             fa_out = np.dot(phi.T, theta)
147         else:
148             phi = self.computeFeatureWeight(inputData)
149             fa_out = np.dot(phi.T, theta)
150         return fa_out

```

4.4 Experimental methods on human

METHODS

Subjects Twenty healthy male volunteers participated in the study. Their average age was $xx.x$ years ($SD = xx.x$ years), height xxx cm ($SD = x.x$ cm) and body mass xx kg ($SD = x.x$ kg). All subjects were right-handed. Exclusion criteria were neurological, vestibular, locomotion, visual disorders and recent limb injuries (self-reported). Prior to their participation, the subjects were informed about the course of the study and were required to sign an informed consent approved by the National medical ethics committee (No. 112/06/13).

Measurement Protocol Subject sat on a chair in front of a TV screen and a 3 axes haptic manipulator (HM) (Haptic master Mk2, MOOG, Nieuw-Vennep, The Netherlands). The TV screen was located 2 m in front of the chair backrest. The subject's right wrist was immobilised and connected to the HM by means of the gimbal mechanism (ADL gimbal mechanism, MOOG, Nieuw-Vennep, The Netherlands). HM constrained the motion of the subjects right hand (hand from here on) to the horizontal plane at the height of the subjects shoulders. The subjects right elbow was suspended from the ceiling by a long string in order to compensate for the effect of gravity and to restrain the motion of the arm to the horizontal plane. In order to fix the position of the shoulders, the trunk of the subject was immobilised by tying it to the backrest of the chair.

The subject was allowed to move the hand from 12 initial circular areas with 10 mm diameter spread in front of the subject to a single target that was located symmetrically to the subjects right shoulder on a virtual wall at a distance equal to 95 (defined as the distance between shoulder and wrist). The initial areas were spread in two lines parallel to the virtual wall. One line was located at the distance equal to 35 wall and included 5 initial areas while the other line was located at the distance equal to 55 included 7 initial areas. The initial areas on each line were placed symmetrically with respect to the target on the virtual wall and were spread with 10 cm distance in-between. The initial areas were spread in the area well within the area of arm motion where the passive joint torques are negligible. The size of the target on the virtual wall was either 5 mm, 10 mm, 20 mm or 40 mm wide. The position of the current initial area, the position of the target and the position of the subjects hand were drawn in 2D perspective on the screen in real time. The size of the initial areas were shown on the screen while the size of the target was hidden and only a cross mark indicated its centre. The experimental setup is depicted in Fig. 1.

The experiment was divided into four parts corresponding to the four randomly selected sizes of the target. In each part, the randomly selected initial position from where the subject had to perform the motion was indicated in red. When the subject reached and stayed within the initial position

for more than 1 s, the colour of the patch changed from red to green. This indicated the subject that the motion can be started. The time of the motion from the moment when the hand left the initial circular patch to the moment when the hand touched the virtual wall was recorded and subtracted from the total time of 90 seconds. The remaining time was displayed on the screen at the bottom of the screen. Success in hitting the target on the virtual wall was clearly indicated by changing the color of the virtual wall from grey to green for 1 s. Besides, the subject obtained a money award of 10 euro cents. If the target was missed, the virtual wall turned red and no reward was given to the subject. The cumulative reward was displayed on top of the screen using large bold fonts.

The task of the subject was to obtain as high reward as possible in the given time. In effect, the subjects had to find the optimal balance between precision, speed of motion and its cost in order to maximise the reward without knowing the actual size of the target. To amplify the effect of cost of motion, HM emulated a viscous media through which the subject had to move the hand.

Bibliography

- [1] R. J. Bootsma, L. Fernandez, and D. Mottet. Behind Fitts law: kinematic patterns in goal-directed movements. *International Journal of Human-Computer Studies*, 61(6):811–821, Dec. 2004.
- [2] R. J. Bootsma, R. G. Marteniuk, C. L. MacKenzie, and F. T. Zaal. The speed-accuracy trade-off in manual prehension: effects of movement amplitude, object size and object width on kinematic characteristics. *Experimental Brain Research*, 98(3):535–41, Jan. 1994.
- [3] E. R. Crossman and P. J. Goodeve. Feedback control of hand-movement and fitts’ law. *The Quarterly Journal of Experimental Psychology*, 35(2):251–278, 1983.
- [4] M. Dean, S.-W. Wu, and L. T. Maloney. Trading off speed and accuracy in rapid, goal-directed movements. *Journal of Vision*, 7(5):1–12, 2007.
- [5] A. A. Faisal, L. P. J. Selen, and D. M. Wolpert. Noise in the nervous system. *Nature Reviews Neuroscience*, 9(4):292–303, 2008.
- [6] P. M. Fitts. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 47(6):381–91, June 1954.
- [7] C. M. Harris and D. M. Wolpert. Signal-dependent noise determines motor planning. *Nature*, 394:780–784, 1998.
- [8] H. Kambara, K. Kim, D. Shin, M. Sato, and Y. Koike. Learning and generation of goal-directed arm reaching from scratch. *Neural networks*, 22(4):348–61, 2009.
- [9] M. Katayama and M. Kawato. Virtual trajectory and stiffness ellipse during multijoint arm movement predicted by neural inverse models. *Biological Cybernetics*, 69(5):353–362, 1993.
- [10] M. Laurent. Fitts’ law in two-dimensional task space. *Experimental Brain Research*, 100:144–148, 1994.

- [11] W. Li. *Optimal control for biological movement systems*. PhD thesis, University of California, San Diego, 2006.
- [12] D. Mitrovic, S. Klanke, and S. Vijayakumar. Adaptive optimal control for redundantly actuated arms. In *Proceedings of the Tenth International Conference on Simulation of Adaptive Behavior*, 2008.
- [13] R. Plamondon and A. M. Alimi. Speed/accuracy trade-offs in target-directed movements. *The Behavioral and brain sciences*, 20(2):279–303; discussion 303–49, June 1997.
- [14] L. Rigoux. *Compromis entre efforts et récompenses : Un modèle unifié de la décision et de la motricité*. Thèse de doctorat, Université Pierre et Marie Curie, 4 place Jussieu 75005 Paris, Nov. 2011.
- [15] L. Rigoux and E. Guigon. A model of reward and effort based optimal decision making and motor coordination. *PLoS Computational Biology*, 2012.
- [16] I. Scott MacKenzie. A note on the information-theoretic basis for fitts law. *Journal of Motor Behavior*, 21(3):323–330, 1989.
- [17] R. Shadmehr, J.-J. Orban de Xivry, M. Xu-Wilson, and T.-Y. Shih. Temporal discounting of reward and the cost of time in motor control. *Journal of Neuroscience*, 30(31):10507–10516, 2010.
- [18] N. Smyrnis, I. Evdokimidis, T. Constantinidis, and G. Kastrinakis. Speed-accuracy trade-off in the performance of pointing movements in different directions in two-dimensional space. *Experimental Brain Research*, 134(1):21–31, Aug. 2000.
- [19] J. F. Soechting. Effect of Target Size on Spatial and Temporal Characteristics of a Pointing Movement in Man. *Experimental Brain Research*, 54:121–132, 1984.
- [20] E. Todorov. Optimality principles in sensorimotor control. *Nature Neurosciences*, 7(9):907–915, 2004.
- [21] E. Todorov. Stochastic Optimal Control and Estimation Methods Adapted to the Noise Characteristics of the Sensorimotor System. *Neural Computation*, 17(5):1084–1108, 2005.
- [22] E. Todorov and M. I. Jordan. Optimal feedback control as a theory of motor coordination. *Nature Neurosciences*, 5(11):1226–1235, 2002.