

# An Energy Efficient Dynamic Gait for a Nao Robot

Zhenglong Sun, Nico Roos

Department of Knowledge Engineering, Maastricht University  
P.O. Box 616, 6200 MD Maastricht, The Netherlands

**Abstract**—This paper presents a framework to generate energy efficient dynamic human-like walk for a Nao humanoid robot. We first extend the inverted pendulum model with the goal of finding an energy efficient and stable walking gait. In this model, we propose a leg control policy which utilizes joint stiffness control. We use policy gradient reinforcement learning to identify the optimal parameters of the new gait for a Nao humanoid robot. We successfully test the control policy in a simulator and on a real Nao robot. The test results show that the new control policy realizes a dynamic walk that is more energy efficient than the standard walk of Nao robot.

**Keywords**—Humanoid Robot, Learning Control, Energy-efficient

## I. INTRODUCTION

The gait of humans is often assumed to be the most energy efficient way of walking [1]. Srinivasan and Ruina [2] confirm this hypothesis using a simple model in which the human is a point mass with straight legs that can change in length during a step. Their results show that the dynamically stable inverted pendulum walk [3] is the most energy efficient gait. The question is whether this result also holds for humanoid robots. In their model, energy consumption is determined by the positive work while increasing the length of the leg. The spring behaviour of human muscles is ignored as well as the energy required to maintain a torque on the knee joint when the leg is not stretched. The torque generated by a bended knee might be ignored while walking. Humans need not bend the knee of the stance leg when walking because they can push off using the foot and the calf muscle. In this way humans can “increase” the length of the leg without torque on the knee joint. A humanoid robot such as a Nao, differs in important ways from a human [4]. First, unlike muscles, motors of robot do not behave like springs. Second, our experiment with a Nao shows that the energy consumption of a motor depends more on the torque than on the work that is done. However, there is no torque on the knee joint of a human when its leg is stretched while the human can still stretch the leg a bit more by pushing off using the foot. Third, a robot such as a Nao cannot push off, because it can not bend its foot. So, the third difference makes it possible to ignore the torque in the human model but not in models of certain robots such as a Nao.

Our main contribution is the development of framework for generating a human-like energy efficient dynamic gait for humanoid robots. A dynamic gait is a gait that is temporarily unstable in the forward direction during a step. In the paper, we first briefly describe *Inverted Pendulum Model* (IPM) [5, 6] which has been adapted to a humanoid robot such as Nao. We extend the IPM with a double support phase which is needed for side-way stability. We used the extended inverted

pendulum model (EIPM) to identify the most energy efficient gait (Section II and III). The results are used to create a gait controller (Section IV) which is fine-tuned for a Nao robot using a *webots* simulator (Section V). The gait is evaluated on a real Nao robot (Section VI). The last section concludes the paper.

## II. THE EXTENDED INVERTED PENDULUM MODEL (EIPM)

The gait of humans and of humanoid (bipedal) robots is a repeating pattern consisting of two phases; a single support phase (SSP) where the body is supported by only one leg and a double support phase (DSP) where the body is supported by both legs [7]. In the DSP the weight of the body is shifted from one leg to the other. The DSP is crucial for the sideways stability but is sometimes ignored when analysing the gait. However, since it is impossible to implement a gait on a Nao without a double support phase, we must consider it in our model. We will start presenting a model without a DSP and subsequently extend the model with a DSP.

The goal of our model is to identify an energy minimal gait for a bipedal robot such as a Nao. To analyse the energy consumption, we need a model of the main joints and limbs of the robot. Since the walk pattern is what we are concerned about, to simplify the model, we assume that the upper part of body above the hip can be ignored. Moreover, since we are interested in a dynamic gait instead of the statically stable gait that is often used [8], we also ignore the feet and the ankle joints. Finally, we assume that legs are weightless, making it possible to use a single point mass for the whole robot. The resulting model consists of 5 links and a point  $m$  shown by the solid line in Fig. 1. We can replace the 5-link model by an equivalent 2-link model consisting of two links that can change their length during a step and a point mass  $m$ , see the dashed lines in Fig. 1. The two links will be denoted as telescopic legs [6, 9, 10].

### A. The model without double support phase

1) *leg-length policy*: If we would know the force produced by the leg, we could apply Newton’s second law to derive the second order equation for the movement of the mass  $m$ . Since the ground generates a reaction force which can become infinitely large at the moment a stretched leg impacts with the ground, and since the length of the leg is bounded by  $l$ , we choose a different approach. We assume a leg-length policy  $\delta(\beta)$  where  $\delta : [-\frac{\pi}{2}, \frac{\pi}{2}] \rightarrow [0, 1]$  and where  $\beta$  is the angle between stance leg with vertical axis and  $l\delta(\beta)$  is the actual length of the leg, as shown in Fig. 1. We use the leg-length policy to determine the radial force  $F_r$  on the leg and a force

$F_t$  perpendicular to  $F_r$ . Note that  $F_t$  works perpendicular to leg while the path of the mass  $m$  need not be perpendicular to the leg because the length of the leg may change. Using the force  $F_t$ , we can determine the movement  $x_t$  perpendicular to leg which is given by:

$$F_t = ma = m \frac{d^2 x_t}{dt^2}$$

$F_t$  is determined by the component of the gravity working perpendicular to the stance leg:  $mg \sin \beta$  and the friction:  $b \frac{dx_t}{dt}$ . Note that we assume the friction to be linear in the speed. The air friction is assumed to be linear because of the low walking speed.

$$\frac{d^2 x_t}{dt^2} + \frac{b}{m} \frac{dx_t}{dt} - g \sin \beta = 0 \quad (1)$$

We can transform this movement in a change of the angle  $\beta$  using:

$$d\beta = \frac{dx_t}{\delta(\beta)l}$$

Therefore,

$$\begin{aligned} \frac{dx_t}{dt} &= \delta(\beta)l \frac{d\beta}{dt} \\ \frac{d^2 x_t}{dt^2} &= l \frac{d\delta(\beta)}{dt} \frac{d\beta}{dt} + \delta(\beta)l \frac{d^2 \beta}{dt^2} \\ &= l \frac{d\delta(\beta)}{d\beta} \left( \frac{d\beta}{dt} \right)^2 + \delta(\beta)l \frac{d^2 \beta}{dt^2} \end{aligned}$$

This result in:

$$\frac{d^2 \beta}{dt^2} + \frac{1}{\delta(\beta)} \frac{d\delta(\beta)}{d\beta} \left( \frac{d\beta}{dt} \right)^2 + \frac{b}{m} \frac{d\beta}{dt} - \frac{g}{\delta(\beta)l} \sin \beta = 0$$

To use this second order differential equation in a simulator, we reformulated it as a system of first order differential equations.

$$\begin{aligned} \frac{d\omega}{dt} &= -\frac{1}{\delta(\beta)} \frac{d\delta(\beta)}{d\beta} \omega^2 - \frac{b}{m} \omega + \frac{g}{\delta(\beta)l} \sin \beta \\ \frac{d\beta}{dt} &= \omega \end{aligned} \quad (2)$$

The length of the stance leg at the beginning and the end of a step, denoted by  $l_1$  and  $l_2$  respectively, need not be the same. We therefore need to know the angle  $\beta$  at the beginning and the end of a step, denoted by  $\beta_1$  and  $\beta_2$  respectively, given a fixed step size  $s$ . The leg-length policy should describe the changes in the leg-length between these angles. To determine the angles  $\beta_1$  and  $\beta_2$ , we apply the cosine rule. Hence,

$$\beta_1 = \arcsin \frac{s^2 + l_1^2 - l_2^2}{2sl_1} \quad \beta_2 = \arcsin \frac{s^2 + l_2^2 - l_1^2}{2sl_2}$$

2) *Force on Stance Leg*: The solution of the above presented system of differential equations enables us to determine the radial force on the stance leg. This force together with the leg-length policy  $\delta(\beta)$  determines the energy consumption of the knee joint of the stance leg. The radial force consists of a gravitational component and a component needed to accelerate the mass in the direction of the radius. The former is a reaction force equal to:  $mg \cos \beta$ , and the latter is determined by the second derivative of the leg-length policy:

$$\begin{aligned} \frac{dl\delta(\beta)}{dt} &= l \frac{d\delta(\beta)}{d\beta} \frac{d\beta}{dt} \\ \frac{d^2 l\delta(\beta)}{dt^2} &= l \left( \frac{d^2 \delta(\beta)}{d\beta^2} \left( \frac{d\beta}{dt} \right)^2 + \frac{d\delta(\beta)}{d\beta} \frac{d^2 \beta}{dt^2} \right) \end{aligned}$$

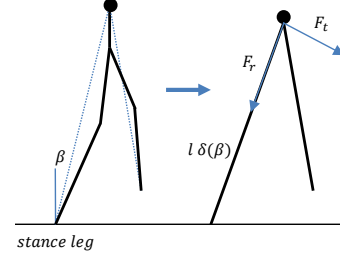


Fig. 1. A Five-links Model Simplified to a Two-links Model

So,

$$F_r = mg \cos \beta + ml \left( \frac{d^2 \delta(\beta)}{d\beta^2} \left( \frac{d\beta}{dt} \right)^2 + \frac{d\delta(\beta)}{d\beta} \frac{d^2 \beta}{dt^2} \right) \quad (3)$$

When the foot of the robot impacts with the ground at the beginning of a step, the direction in which the mass  $m$  is moving may change. Since the change in direction is instantaneous, no energy is transferred to the mass. The conservation law of kinetic energy now implies that the speed of the mass may not change the moment the foot impacts with the ground. For a gait without acceleration, this implies that the speed of  $m$  at the beginning and end of a step must be the same.

Also the law of momentum conservation applies. The sum of the momentum before and after the foot impacts with the ground must be 0. This does not imply that the impulse generated by the reaction force of the leg when it impacts with the ground is 0. This impulse causes the change in the direction of  $m$ . The impulse generated by the reaction force of the leg is determined by the change in speed of  $m$  in the direction of the *new* stance leg:

$$I_r = v_{r,b}m - v_{r,e}m$$

The subscript  $r$  denotes the radial direction of the leg, the subscript  $b$  denotes the beginning of a step and the subscript  $e$  the end. After the foot impacts with the ground, the speed in the direction of the leg is:

$$v_{r,b} = \frac{dl\delta(\beta)}{dt}(t_b)$$

Before the foot impacts with the ground we have to calculate the component of the speed in the direction of the new stance leg. We first calculate the speed in the  $x$  and  $z$  direction of the Cartesian coordinate system.

$$\begin{aligned} v_t &= l\delta(\beta) \frac{d\beta}{dt} & v_r &= \frac{dl\delta(\beta)}{dt} = l \frac{d\delta(\beta)}{d\beta} \frac{d\beta}{dt} \\ v_x &= v_t \cos(\beta) + v_r \sin(\beta) & v_z &= v_t \sin(\beta) + v_r \cos(\beta) \end{aligned}$$

So, impulse produced by the leg becomes:

$$I_r = m(v_{r,b} - (v_{x,e} \sin(\beta_b) + v_{z,e} \cos(\beta_b)))$$

The impulse is also equal to:  $I_r = \int F_i dt$ , which enables us to calculate the force  $F_i$  generated by the impulse. In an ideal situation the impact time with the ground is infinitely small implying an infinitely large reaction force  $F_i$  produced

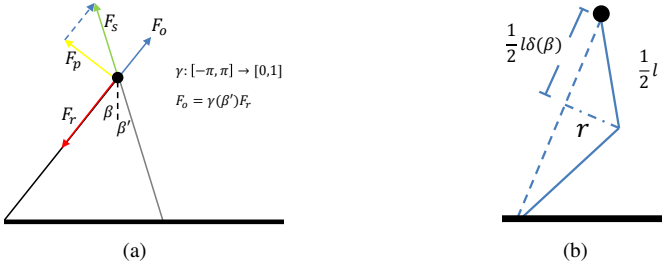


Fig. 2. (a) Force Policy in Double Support Phase and (b) Arm of Knee Bending

by the leg on the mass  $m$ . In practice material always bends or compresses somewhat. This increases the impact time and thereby reduces the reaction force of the leg. Although the force will now be finite, even if we would know the impact time, we cannot calculate it. We therefore make the simplifying assumption that the reaction force is constant during the impact. Moreover, we assume that the impact time is 10 ms. These assumptions make it possible to calculate the force  $F_i$  on the leg during the impact time.  $F_i = 0$  outside the impact time of 10 ms.

### B. The model with double support phase

We extended the model described in the previous subsection with double support phase. We stop the single support phase at 75% of a step [11]; i.e., at the angle  $\beta_1 + 0.75(\beta_2 - \beta_1)$  where  $\beta_1$  and  $\beta_2$  are the begin and the end angle, respectively, of the stance leg during a step. From this moment the swing leg will also be on the ground, thereby influencing the movement of the mass  $m$ .

We need a way to describe the influence of the swing leg on the mass in DSP. This can not be done by just simply applying a leg-length policy for the swing leg in the double support phase. Given the step size  $s$  and the leg-length policy  $\delta(\beta)$  of the stance leg, the length leg of the swing leg is fixed. Prescribing this length of the swing leg by a policy, creates a rigid triangle in which the mass  $m$  can no longer move freely. We therefore choose to let the mass  $m$  move freely given the leg-length policy of the stance leg and use a force policy for the swing leg in the double support phase. So the length of the swing leg is determined by the leg-length policy of the stance leg and the step size, but force the swing leg executes on the mass is determined by the force policy of the swing leg. This force may influence the forward speed of the mass  $m$ .

The force  $F_s$  generated by the swing leg, causes a force  $F_o$  in the opposite direction of  $F_r$  and a force  $F_p$  perpendicular to the swing leg (as shown in Fig. 2(a)). We use the force policy  $\gamma(\beta')$  where  $\gamma : [-\frac{\pi}{2}, \frac{\pi}{2}] \rightarrow [0, 1]$  and  $\beta'$  is the angle of the swing leg with the vertical axis, to determine  $F_o = \gamma(\beta')F_r$ , and thereby the force  $F_p$  and the force  $F_s$  generated by the swing leg. Note that the force policy  $\gamma(\beta')$  ensures that the stance leg stays on the ground. Also note that the angle  $\beta$  of the stance leg with the vertical axis, the leg-length policy  $\delta(\beta)$ , and the step size, determines the angle  $\beta'$  of the swing leg with the vertical axis.

We can now derive the following differential equations for the movement of the mass  $m$  in the DSP:

$$\frac{d^2\beta}{dt^2} + \frac{1}{\delta(\beta)} \frac{d\delta(\beta)}{d\beta} \left( \frac{d\beta}{dt} \right)^2 + \frac{b}{m} \frac{d\beta}{dt} - \frac{g}{\delta(\beta)l} \sin \beta + F_r \gamma(\beta') \tan(\beta + \beta') = 0 \quad (4)$$

The last term of Equation 4 is determined by the force  $F_p$ . The force  $F_r$  in this term is given in Equation 3.

$$F_p = F_r \gamma(\beta') \tan(\beta + \beta')$$

### C. The energy consumption

To calculate the energy consumption of the robot, we make use of the fact that the robot has to bend its leg at the knee joint in order to shorten the leg. The energy consumption is assumed to be proportional to the torque of these joints. So, a stretched leg requires no energy while an almost completely bended leg requires a maximum amount of energy. The experiment with a Nao described in Section III, confirmed our assumption that the torque on the knee joint determines its energy consumption. We will use this observation to determine the energy consumption in the model.

The torque on the knee joint is determined by the force on the stance leg, and the arm  $r$ , shown in Figure 2(b). The arm  $r = \frac{1}{2}l\sqrt{1 - \delta(\beta)^2}$  is a function of the shortening of the leg specified by the leg-length policy  $\delta(\beta)$ . The radial force on the stance leg  $F_{rt}$  is determined by  $F_r$  defined in Equation 3,  $F_o$  and by the impulse when the swing leg hits the ground. That is

$$F_{rt} = F_r + F_i - F_o$$

We determine the energy consumption based the observation that the energy consumption is linear increasing with the torque. Therefore, we define the energy consumption as:

$$E = \int_{-\frac{1}{2}T}^{\frac{1}{2}T} \frac{1}{2}l\sqrt{1 - \delta(\beta)^2} F_{rt} dt$$

In this equation, the shortening of the leg  $\delta(\beta) \in [0, 1]$  is the policy that is used to control the gait. This means that the bending angle of the knee-joints and the time that the knee-joints are bended are key factors in determining the energy consumption.

## III. SIMULATION EXPERIMENTS WITH THE EIPM

This section describes three experiments. The first experiment confirms our assumption that knee joint provides the main contribution to the energy consumption of a walk. The second and third experiment describe the identification of the optimal policy.

### A. Torque

As mentioned in the introduction, we assume that the energy consumption of a gait is determined by the torque on the knee joint. We can ignore the torque on the ankle joint, because we assume a dynamic gait. Since the upper part of the body is balanced above the hip joint, ideally, no torque is required in the hip joint. Experiment on the Nao by Kulk and Welsh [12] confirms the assumption about the ankle and hip joints. Moreover they show that the knee joint provides the main contribution to the energy consumption of a walk.

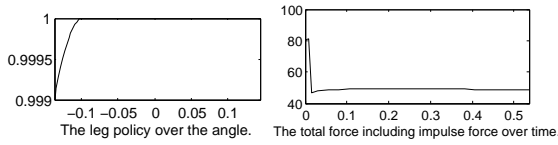


Fig. 3. The "energy optimal" leg-length policy

In the first experiment, we measure the electric current of a knee joint. Since the voltage is constant, the energy consumption is linear in the current. We changed the position of the Nao from standing upright (knee angle =  $180^\circ$ ) to a position where the angle is around  $90^\circ$ . We kept the Nao for a while in this position after which we let the Nao move back to the upright position. During this exercise, we monitored the current of a knee joint. When standing upright it is around 0.2 A, and decreases to 0.1 A during the bending of the knees. In the fixed bended position, the current is 0.7 A. Finally, when the Nao moves back to the upright position, the current gradually drops from 0.7 A to 0.2 A. During these experiment "Smart Stiffness" was switched on. Based on this experiment, we conclude that the torque determine energy consumption and that the positive work can be ignored.

#### B. Optimal leg-length policy in the absence of a double support phase

In the first series of simulation experiments, we generated a large number of leg-length policies. We divided a step into eight intervals and chose a leg-length for the beginning (end) of each interval. Next we determined a polynomial for the leg-length policy through these points using the cubic spline method. We repeated the process to investigate all possible policies with a resolution of eight intervals, to find an optimal policy that minimizes the power usage. We also considered deviations in other numbers of intervals. Eight intervals gave the best results with respect to accuracy and computation time.

Fig. 3 shows that a slightly bended leg which is subsequently stretched to the maximum length results in the most energy-efficient gait. The optimal policy keeps the leg stretched during the remainder of the step. The energy consumption of this policy for one step is 0.011. Fig. 3 also shows a relatively large peak force on the leg, which results from the impulse when leg hits the ground.

#### C. Optimal leg-length and force policy in the presence of a double support phase

In a second series of simulation experiments, we added the double support phase and divided the double support phase in to four intervals and chose a force percentage of the beginning (end) of each interval. Next we determined a polynomial for the force policy through these points using the cubic spline method. We iterate the process to investigate all possible combinations of leg-length and force policies to find an optimal combination that minimizes the power usage.

The experiments showed that the optimal leg-length policy is the same as the leg-length policy for the gait without DSP, as shown in Fig. 3. The optimal force policy is the policy that sets the force generated by the swing leg to zero till the end of the DSP. This implies that the force generated by the swing

leg in the DSP will only be determined by requirement of the sideways stability.

## IV. GAIT CONTROLLER AND PARAMETERS

This section presents the parameters of a gait that realizes the leg-length policy determined by the experiments described in the previous section. Moreover, the stiffness control of the joints is discussed.

### A. Gait Parameters

Based on the results of the simulation experiments, we identify six parameters that are essential in controlling a dynamic gait:

- *Step Size* ( $\theta_1$ ): Defines the how long Nao can move in a single step (forward).
- *Ankle Roll* ( $\theta_2$ ): Defines the roll angel of ankle joints.
- *Step Height* ( $\theta_3$ ): Defines the maximum distance between ground and lifting feet. A high step height will require higher speed of the swing leg and may cause horizontal instability. A low step height increases the possibility of tripping and limits the step size.
- *Angle Bending* ( $\theta_4$ ): Defines the maximum bending of the new stance leg at the beginning of the single support phase. This parameter determines the forward velocity and the energy cost.
- *Swing Time* ( $\theta_5$ ): Defines how long the single support phase lasts. This parameter determines the forward walking velocity.
- *Stretch Time* ( $\theta_6$ ): Defines how long it takes for the stance leg to stretch to its full length at the beginning of the single support phase.
- *Torso Pitch Inclination* ( $\theta_7$ ): Defines the maximum angle that torso leans in forward direction. If positive, it will move the center of mass (COM) forward. If it is set not appropriate, a fall will occur.

TABLE I. TRAJECTORY PARAMETERS

Description of joint motion	q
step size	$\theta_1$
swing hip pitch	$p_1(\theta_1, \theta_3, \theta_5, \theta_7)$
swing hip roll	$\theta_2$
swing knee pitch	$p_2(\theta_1, \theta_3, \theta_5, \theta_7)$
swing ankle pitch	$-p_1 - p_2$
swing ankle roll	$-\theta_2$
stance hip pitch	$p_3(\theta_1, \theta_3, \theta_4, \theta_6, \theta_7)$
stance hip roll	$\theta_2$
stance knee pitch	$p_4(\theta_1, \theta_3, \theta_4, \theta_6, \theta_7)$
stance ankle pitch	$-p_3 - p_4$
stance ankle roll	$-\theta_2$

Table I shows all the parameters of the trajectory for walk movement. Walk posture  $q$  is determined by joints value, step size, acceleration and so on. The value of hip pitch, knee pitch and ankle pitch are functions  $p_n$  of parameter sets.

### B. State-dependent stiffness

An appropriate joint stiffness of the knee and ankle passively safeguard the internal leg stability and control the energy consumption [13, 14]. Full joints stiffness is good for manipulating the movement of a robot, but may lead to high energy consumption and the absence of a dynamic gait.

The two-link EIPM does not model the foot, therefore no torque will be applied on the ankle joints. So when design the gait controller in the simulator and the real Nao, we set the stiffness of stance foot ankle to *zero* in the single support phase. This ensures that the stance leg can move in exactly the same way as we assumed in the EIPM. Since there is no torque on the ankle, there is no energy cost.

We need a sufficiently high stiffness in the ankle of swing leg to keep the foot sole be parallel to ground when swing foot is about to touch the ground. Due to unpredictable small body movements, foot sole may not always be parallel to the ground. This could result in "toe poking" or "heel striking". Therefore, we lower the stiffness of the swing leg ankle to *zero* just before the foot touches the ground.

## V. OPTIMIZING GAIT PARAMETER FOR NAO

This section describes the optimization of the control parameter of a dynamic gait for a Nao.

### A. Policy Gradient Algorithm

After investigated several policy search algorithms [15, 16], we chose to use a policy gradient algorithm presented by Kohl and Stone [17] to optimize the Nao's gait. The objective function  $F$  that is optimized by the algorithm, will be a function of the energy cost and the stability of the Nao. The policy gradient algorithm starts with an initial parameter vector  $\pi = \theta_1, \dots, \theta_N$  and estimates the partial derivative of the objective function  $F$  with respect to each parameter. This is done by evaluating  $t$  randomly generated policies  $R_1, \dots, R_t$  near  $\pi$ , such that each  $R_i = \theta_1 + \delta_1, \dots, \theta_N + \delta_N$  and  $\delta_j$  is randomly chosen to be either  $-\epsilon$ ,  $0$ ,  $+\epsilon$ , where  $\epsilon$  is a small fixed value relative to  $\theta$ . After evaluating each policy  $R_i$  on the objective function  $F$ , each dimension of every  $R_i$  is grouped into one of the three categories to estimate an average gradient for each dimension:

$$R_i \in \begin{cases} S_{+\epsilon,n}, & \text{if the } n\text{th parameter of } R_i \text{ is } \theta_n + \epsilon_n \\ S_{+0,n}, & \text{if the } n\text{th parameter of } R_i \text{ is } \theta_n + 0 \\ S_{-\epsilon,n}, & \text{if the } n\text{th parameter of } R_i \text{ is } \theta_n - \epsilon_n \end{cases}$$

We calculate average score  $Avg_{-\epsilon,n}$ ,  $Avg_{+0,n}$  and  $Avg_{+\epsilon,n}$  for  $S_{-\epsilon,n}$ ,  $S_{+0,n}$  and  $S_{+\epsilon,n}$  respectively.

- $Avg_{-\epsilon,n}$  The average score for all  $R_i$  that have a negative perturbation in dimension  $n$
- $Avg_{+0,n}$  The average score for all  $R_i$  that have a zero perturbation in dimension  $n$
- $Avg_{+\epsilon,n}$  The average score for all  $R_i$  that have a positive perturbation in dimension  $n$

These three average values estimate the benefit of altering the  $n$ th parameter by  $+\epsilon_n$ ,  $0$ ,  $-\epsilon_n$ . An adjustment vector  $A$  of size  $n$  is calculated where:

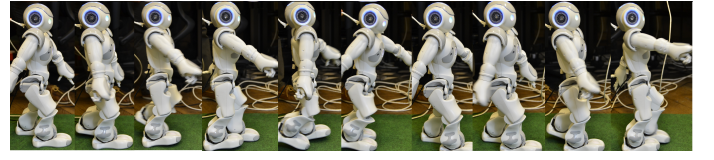


Fig. 4. Series of snapshots of best learned dynamic walking with our method

$$A_n = \begin{cases} 0, & \text{if } Avg_{+0,n} \geq Avg_{+\epsilon,n} \text{ and } Avg_{+0,n} \geq Avg_{-\epsilon,n} \\ Avg_{+\epsilon,n} - Avg_{-\epsilon,n}, & \text{otherwise} \end{cases}$$

---

### Algorithm 1 Pseudo-code of Policy Gradient Algorithm

---

```

 $\pi \leftarrow \text{InitialPolicy}$ 
while !done do
   $R_1, R_2, \dots, R_t = t \text{ random perturbations of } \pi$ 
   $\text{evaluate}(R_1, R_2, \dots, R_t)$ 
  for  $n = 1$  to  $N$  do
     $Avg_{+\epsilon,n}$ 
     $Avg_{+0,n}$ 
     $Avg_{-\epsilon,n}$ 
    if  $Avg_{+0,n} > Avg_{+\epsilon,n}$  and  $Avg_{+0,n} > Avg_{-\epsilon,n}$ 
    then
       $A_n \leftarrow 0$ 
    else
       $A_n \leftarrow Avg_{+\epsilon,n} - Avg_{-\epsilon,n}$ 
    end if
  end for
   $A \leftarrow \frac{A}{|A|} * \eta$ 
   $\pi \leftarrow \pi + A$ 
end while

```

---

In order to generate a gait that is energy efficient and stable, we adopt an objective function based on the energy cost and the stability. The energy cost is expressed by the normalized current  $M_c$ , and the stability by normalized standard deviation of the three accelerometers  $M_a$

$$F = 1 - (w_c M_c + w_a M_a) \quad (5)$$

The components of the objective function are weighted by  $w_c$  and  $w_a$  respectively to emphasize the relevance of the underlying goals. These weights are constrained so that the sum of the weights are equal to one. In this experiment, we set  $w_c = 0.25$ ,  $w_a = 0.75$ .

### B. Learning optimal parameters in simulator

Experiments were conducted to generate the optimal gait parameters and to test the gait's performance. Modules for the dynamic gait and for the policy gradient algorithm were uploaded in Nao model of the *Webots* simulator. We used a relatively elementary hand-tune gait as a starting policy for the policy gradient algorithm described in Section V. Though a bad starting policy may lead to a simulation failure, we did not deliberately optimize the starting policy. The performance of the task is measured by monitoring the knee torque and electric current. Falling or "toddle" is penalized. We found



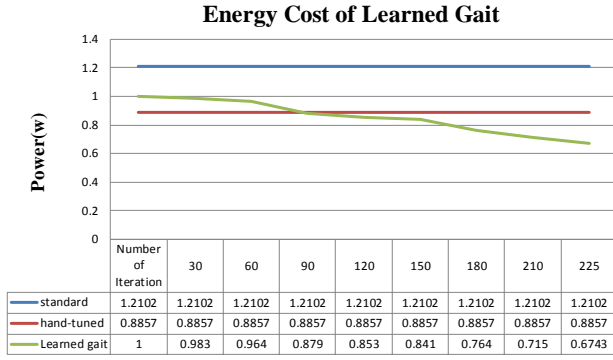


Fig. 5. Comparison of Energy Cost of Three Models

a most energy-efficient gait after 237 iterations. Subsequent evaluations showed no further improvement. The learning algorithm produce a set parameters of a stable gait which is more energy-efficient and faster than the standard walk of Nao. The parameters of the gait and  $\epsilon$  value are given in Table II.

TABLE II. INITIAL PARAMETERS AND BEST LEARNED PARAMETERS

Parameter	Initial Value	epsilon	Learned Value
step size	6(cm)		6(cm)
step height	3(cm)	0.02	3.24(cm)
angle bending	15 (degree)	0.1	13.8 (degree)
swing time	300(ms)	25	225(ms)
bending time	300(ms)	25	375(ms)
torso pitch inclination	10 (degree)	0.1	7.5 (degree)

## VI. SIMULATION AND REAL WORLD EVALUATION

To validate our approach, we perform a real world experiment with the Nao V3.2 humanoid robot, which has 25 degrees of freedom. We evaluated the efficacy of the learned parameters by sending them to a Nao robot and command the Nao to walk a constant distance. We compared the energy consumption of learned gait with the energy consumption of the standard gait of the Nao and our initial hand-tuned gait. The step size was set to 6 cm. Fig. 5 shows the energy consumption of the three gaits. We see that the learned gait results in a power reduction of 44% of standard gait and a reduction of 23.8% of hand-tune gait.

## VII. CONCLUSION

In this paper, we presented a framework to generate energy efficient dynamic human-like walk for a Nao humanoid robot. We extended the inverted pendulum model and used it to identify a control policy for an energy efficient gait. We optimize the control policy for a Nao humanoid robot and evaluated the result on the real Nao. The latter evaluation shows that our dynamic gait is more energy efficient than the standard gait of the Nao.

## REFERENCES

- [1] S. H. Collins and A. Ruina, "A bipedal walking robot with efficient and human-like gait," in *ICRA*. IEEE, 2005, pp. 1983–1988.
- [2] M. Srinivasan and A. Ruina, "Computer optimization of a minimal biped model discovers walking and running," *Nature*, vol. 439, pp. 72–75, 2006.

- [3] A. D. Kuo, "The six determinants of gait and the inverted pendulum analogy: A dynamic walking perspective," *Human movement science*, vol. 26, no. 4, pp. 617–656, 2007.
- [4] Z. Sun and N. Roos, "An energy efficient gait for a nao robot," *BNAIC*, 2013.
- [5] S. Kajita and K. Tani, "Study of dynamic biped locomotion on rugged terrain-derivation and application of the linear inverted pendulum mode," in *Robotics and Automation, 1991. Proceedings., 1991 IEEE International Conference on*. IEEE, 1991, pp. 1405–1411.
- [6] S. Kajita, O. Matsumoto, and M. Saigo, "Real-time 3d walking pattern generation for a biped robot with telescopic legs," in *Robotics and Automation, 2001. Proceedings 2001 ICRA. IEEE International Conference on*, vol. 3. IEEE, 2001, pp. 2299–2306.
- [7] A. Hamon, Y. Aoustin *et al.*, "Optimal walking gait with double support, simple support and impact for a bipedal robot equipped of four-bar knees," *IMSD 2012*, 2012.
- [8] A. Goswami, "Postural stability of biped robots and the foot-rotation indicator (fri) point," *The International Journal of Robotics Research*, vol. 18, no. 6, pp. 523–533, 1999.
- [9] H. Dong, M. Zhao, and N. Zhang, "High-speed and energy-efficient biped locomotion based on virtual slope walking," *Autonomous Robots*, vol. 30, no. 2, pp. 199–216, 2011.
- [10] M. Zhao, H. Dong, and N. Zhang, "The instantaneous leg extension model of virtual slope walking," in *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on*. IEEE, 2009, pp. 3220–3225.
- [11] J. A. DeLisa, *Gait analysis in the science of rehabilitation*. Diane Publishing, 1998, vol. 2.
- [12] J. Kulk and J. Welsh, "A low power walk for the nao robot," in *Proceedings of the 2008 Australasian Conference on Robotics Automation*, 2008.
- [13] M. Günther and R. Blickhan, "Joint stiffness of the ankle and the knee in running," *Journal of biomechanics*, vol. 35, no. 11, pp. 1459–1474, 2002.
- [14] B. Lim, J. Lee, J. Kim, M. Lee, H. Kwak, S. Kwon, H. Lee, W. Kwon, and K. Roh, "Optimal gait primitives for dynamic bipedal locomotion," in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*. IEEE, 2012, pp. 4013–4018.
- [15] M. P. Deisenroth, R. Calandra, A. Seyfarth, and J. Peters, "Toward fast policy search for learning legged locomotion," in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*. IEEE, 2012, pp. 1787–1792.
- [16] J. Peters and S. Schaal, "Policy gradient methods for robotics," in *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on*. IEEE, 2006, pp. 2219–2225.
- [17] N. Kohl and P. Stone, "Machine learning for fast quadrupedal locomotion," in *AAAI*, vol. 4, 2004, pp. 611–616.