BULGARIAN ACADEMY OF SCIENCES

CYBERNETICS AND INFORMATION TECHNOLOGIES • Volume **12**, No 3

Sofia • 2012

Optimization of a Compact Model for the Compliant Humanoid Robot COMAN Using Reinforcement Learning

Luca Colasanto, Petar Kormushev, Nikolaos Tsagarakis, Darwin G. Caldwell

Department of Advanced Robotics, Istituto Italiano di Tecnologia (IIT), Via Morego 30, 1616 Genova
Emails: { luca.colasanto, petar.kormushev, nikos.tsagarakis, darwin.caldwell } @iit.it

Abstract: *COMAN is a compliant humanoid robot. The introduction of passive compliance in some of its joints affects the dynamics of the whole system. Unlike traditional stiff robots, there is a deflection of the joint angle with respect to the desired one whenever an external torque is applied. Following a bottom up approach, the dynamic equations of the joints are defined first. Then, a new model which combines the inverted pendulum approach with a three-dimensional (Cartesian) compliant model at the level of the center of mass is proposed. This compact model is based on some assumptions that reduce the complexity but at the same time affect the precision. To address this problem, additional parameters are inserted in the model equation and an optimization procedure is performed using reinforcement learning. The optimized model is experimentally validated on the COMAN robot using several ZMP-based walking gaits.*

Keywords: *Humanoid robot, Reinforcement learning, Dynamic walking.*

1. Introduction

The majority of the existing humanoid robots are powered by stiff actuation systems as in Asimo, HRP-3, iCub, LOLA and Hubo [1-11]. In fact, the predominant approach consists of using non-backdrivable, stiff transmission systems and high-gain PID controllers. This solution provides high-precision and high-load disturbance rejection but at the same time it makes the robot unsafe during interaction with humans as the environment. Moreover, the performance in

terms of energy efficiency, peak power limit and overall adaptability to the environment is very limited compared to the human being.

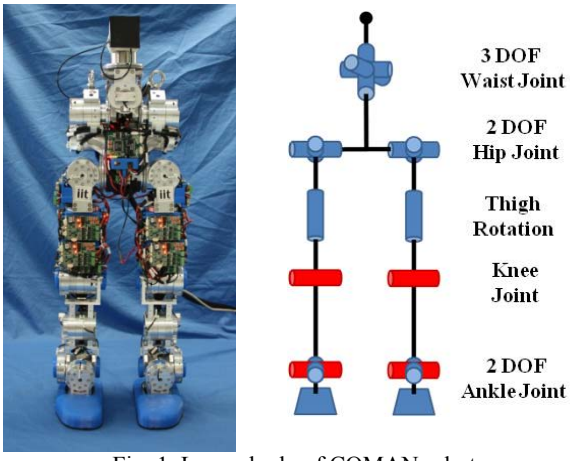

Fig. 1. Lower body of COMAN robot

Following a bioinspired approach, the new COmpliant huMANoid (COMAN) robot (Fig. 1) has been built implementing physical compliance inside the actuation system [12]. In details, passive elastic mechanism is inserted in some joints of the robot (knees and ankles) between the motor and the link. The elastic transmission gives many improvements to the robot during walking reducing the effect of foot impact with the ground. At the same time it adds extra dynamics to the system that is not presented in the common stiff robot.

In this study, a reduced model able to represent the motion of the robot including the effect of compliance is presented, then a learning technique is used to improve the performance of the model. The presentation of the work is organized as follows: In Section 2, the working principle of the model as well as the main equations of the model are presented, in Sections 3 and 4, the model parameters and the proposed optimization approach are reported.

2. Compliant humanoid model

The COMAN robot is a multi degree-of-freedom (dof), non-linear, spring-mass system because of the introduction of passive compliance in the joints. In this section, the modeling procedure to obtain a compact model of the system is reported. Following a bottom-up approach, joint dynamics is identified and then the resultant effects of all the joints are modeled using a Cartesian spring-mass-damper model at the level of the Center of Mass (CoM) [14].

2.1. Joint model

In COMAN, there are two types of joints. In the first type, the motor actuates the link through a harmonic reduction drive group. These joints are called “stiff” joints because the stiffness, due to the harmonic gearbox K_{h}, is very high. “Compliant” joints are the second type of joints where an additional physical elasticity is

incorporated in the actuation. In particular, a passive elastic mechanism is inserted in these joints between the electrical motor and the link. The additional elastic mechanism is in series with the harmonic drive and is characterized by stiffness K_s.

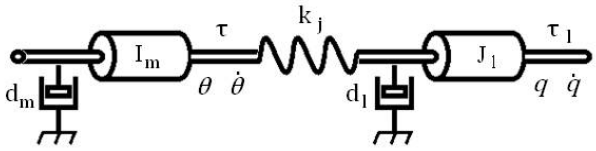

Fig. 2. Joint model

The schematic model of the joint is shown in Fig. 2. Adapting the model from [13] to this representation, the joint can be described by the following equations:

(1) $$I_\mathrm{m}\ddot{\theta} + d_\mathrm{m}\dot{\theta} = \tau - k_j(\theta - q),$$

(2) $$J_\mathrm{l}\ddot{q} + d_\mathrm{l}\dot{q} = \tau_l + k_j(\theta - q),$$

where θ, $\dot{\theta}$ and τ are the position, velocity and torque of the motor, respectively, reflected at the link side after the gear reduction:

(3) $$\theta = \frac{1}{N}\theta_\mathrm{m},$$

(4) $$\tau = N\tau_\mathrm{m},$$

where N is the gear ratio (N=100:1), θ_m and τ_m are the position and torque of the motor; I_m and d_m are the inertia and damping of the motor reflected to the link side as follows:

(5) $$I_\mathrm{m} = N^2 J_\mathrm{m},$$

(6) $$d_\mathrm{m} = N^2\left(D_\mathrm{m} + \frac{K_\tau K_\mathrm{bemf}}{R_\mathrm{m}}\right),$$

where K_τ and K_bemf are the torque sensitivity and back EMF constant, R_m is the stator resistance and D_m is the physical damping of the motor. Finally, q, $\dot{q}$, τ_l, J_l and d_l are the position, velocity, torque, inertia and damping of the link respectively and k_j is the resultant joint stiffness ($k_j = K_\mathrm{h}$ for the stiff joints and $k_j = \frac{K_\mathrm{h}K_\mathrm{s}}{K_\mathrm{h}+K_\mathrm{s}}$ in the case of compliant joints). In the case of the compliant joints, it is possible to approximate the resultant joint stiffness with K_s since $K_\mathrm{h} \approx 8000\ (\mathrm{N.m})/\mathrm{rad}$ is much larger than $K_\mathrm{s} \approx 100\ (\mathrm{N.m})/\mathrm{rad}$. From the stiffness and damping value of each joint it is possible to define the joint stiffness and damping matrix as: $K_\mathrm{j} = \mathrm{diag}(k_j^i)$ and $D_\mathrm{j} = \mathrm{diag}(d_l^i)$, i={1, 6}. Both of them are 6×6 diagonal positive definite matrices.

2.2. Cartesian model at the CoM

Based on the compliant joint model introduced in the previous section, the compliant robot behaviour is approximated by an equivalent Cartesian spring-mass-damper model at the level of the CoM.

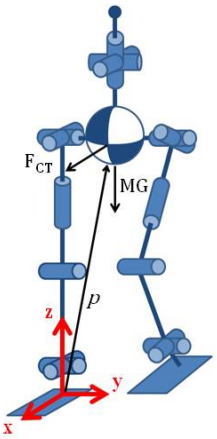

Fig. 3. Robot model and associated support foot reference frame

For each leg, the Jacobian matrix J_{CoM} from the foot base frame placed below the ankle, to the frame placed at the CoM of the robot has been computed (Fig. 3). Using the joint stiffness matrix K_{j}, the resultant Cartesian stiffness matrix $K_{\mathrm{C}} \in R^{6\times6}$ at the pelvis level (CoM) can be obtained by the following equation:

(7) $$K_{\mathrm{C}} = J_{\mathrm{CoM}}^{-\mathrm{T}} K_{\mathrm{j}} J_{\mathrm{CoM}}^{-1}$$

where $J_{\mathrm{CoM}}^{-\mathrm{T}}$ is the inverse transposed Jacobian matrix. In a similar manner, the resultant Cartesian damping matrix $D_{\mathrm{C}} \in R^{6\times6}$ at the pelvis level (CoM) can be obtained as follows:

(8) $$D_{\mathrm{C}} = J_{\mathrm{CoM}}^{-\mathrm{T}} D_{\mathrm{j}} J_{\mathrm{CoM}}^{-1}$$

where D_{j} is the joint stiffness matrix defined in the previous section. Equations (7) and (8) are an approximation of the complete relationship due to the fact that they do not take into account the change of the Jacobian matrix during the deflection movement [15]. Consequently, the approximation is valid when the deflection is small which occurs during the model experiments.

2.3. Working principle of the model

The dynamics model of COMAN is developed with the following assumptions in mind:

(A1) The joints' positions θ are controlled with a stiff PID loop.

(A2) The elasticity in the joint transmission system is due to the harmonic drive compliance as well due to additional physical elasticity integrated in the knee and ankle pitch joints of the leg.

(A3) A single mass approximation is used for the robot model.

The first assumption allows the reduction of the model's complexity. In fact, in the case of an ideally stiff position control (motor position tracking error equal to zero) and high reduction ratio (back-drivability approximately zero), the dynamics of the motor in (1) can be ignored when the robot is subject to external force

perturbations. In this case, (2) approximates the overall joint/link dynamic because the dynamics of the controlled actuator is much faster than the dynamics of the transmission.

During walking experiments, the position deflection of the CoM with respect to the desired position is large along the directions *x* and *z* and smaller along the lateral direction *y* according to foot frame (Fig. 3). This is a consequence of (A2), in fact, the level of compliance is high in the sagittal plane of the humanoid robot (due to additional elasticity in the knee and ankle pitch joints) while in lateral direction the robot is stiffer (only the compliance of the harmonic reduction drive contributes to this). Because of that, in *y* direction the movement can be approximated by a stiff system.

Finally, (A3) is an approach which has been extensively used in trajectory generation and control of humanoid robots [16]. Therefore, according to this assumption, the dynamics of the robot is approximated to the dynamics of the single mass placed at the pelvis (CoM position).

According to the previous consideration and considering equations (7) and (8), the forces generated at the pelvis (CoM) frame when the CoM position $\boldsymbol{p} = [x\ y\ z]^{\mathrm{T}} \in R^3$ deflects with respect to its reference position vector $\boldsymbol{p}_{\mathbf{r}=}[x_{\mathrm{r}}\ y_{\mathrm{r}}\ z_{\mathrm{r}}]^{\mathrm{T}} \in R^3$, can be expressed as follows:

$$F_{\mathrm{CT}} = -K_{\mathrm{CT}}(p - p_{\mathrm{r}}) - D_{\mathrm{CT}}\dot{p} \tag{9}$$

where K_{CT}, $D_{\mathrm{CT}} \in R^{3\times3}$ are sub-matrices of K_{C}, D_{C} related to the linear motion along *x*, *y* and *z*. In case of diagonal matrices *x*, *y* and *z* dynamics are completely decoupled but in this case the off-diagonal elements of the matrices in (7) and (8) are different from zero. Therefore, decoupling the movement of the robot is not possible.

The linear passive dynamics of the single-mass model can be described by the following expression:

$$M\ddot{p} = F_{\mathrm{CT}}(p, \dot{p}) + MG_{\mathrm{CT}} \tag{10}$$

where the mass-matrix $M = \mathrm{diag}(m, m, m) \in R^{3\times3}$ with m being the total mass of the robot placed at the CoM, $F_{\mathrm{CT}} = [f_{cx}\ f_{cy}\ f_{cz}]^{\mathrm{T}} \in R^3$ is the Cartesian forces given by (9) and $G_{\mathrm{CT}} = [0\ 0\ -g]^{\mathrm{T}} \in R^3$ represents the gravity.

Considering only the passive dynamics along the sagittal and vertical directions and referring to equation (9), (10) can be written in a matrix form as follows

$$\begin{bmatrix} m\ddot{x} \\ m\ddot{z} \end{bmatrix} = -\begin{bmatrix} K_{xx} & K_{xz} \\ K_{zx} & K_{zz} \end{bmatrix}\begin{bmatrix} x - x_r \\ z - z_r \end{bmatrix} - \begin{bmatrix} D_{xx} & D_{xz} \\ D_{zx} & D_{zz} \end{bmatrix}\begin{bmatrix} \dot{x} \\ \dot{z} \end{bmatrix} + \begin{bmatrix} 0 \\ -mg \end{bmatrix} \tag{11}$$

where K_{xx}, K_{xz}, K_{zx} and K_{zz} are the relevant elements of K_{C}; D_{xx}, D_{xz}, D_{zx}, D_{zz} are the relevant elements of D_{C}; x_r and z_r are Cartesian position reference of the CoM; and x, z, $\dot{x}$ $\dot{z}$, $\ddot{x}$, $\ddot{z}$ are position, velocity and acceleration of the CoM when it is subject to external loads. The passive dynamic of the CoM of the robot in Cartesian space during stance phase is described by (11).

2.4. Reference trajectory generation

The reference trajectories used later in the learning technique and in the experimental evaluation of the model were generated based on the ZMP approach. The desired gait can be defined by a minimum set of data: step length (sl), single support duration (T_{ss}), double support duration (T_{db}). During the single support the robot was approximated with a single mass linear inverted pendulum as in [16], [17] The ZMP trajectory was computed in order to achieve the desired gait. The reference position of CoM p_{r} can be obtained from the defined ZMP reference $p_{\mathrm{ZMP}} = [x_{\mathrm{ZMP}}\ y_{\mathrm{ZMP}}\ 0]^{\mathrm{T}} \in R^3$.

$$\ddot{x_{\mathrm{r}}} - \frac{g}{Z_{\mathrm{c}}} x_{\mathrm{r}} = -\frac{g}{Z_{\mathrm{c}}} x_{\mathrm{ZMP}}, \tag{12}$$

$$\ddot{y_{\mathrm{r}}} - \frac{g}{Z_{\mathrm{c}}} y_{\mathrm{r}} = -\frac{g}{Z_{\mathrm{c}}} y_{\mathrm{ZMP}} \tag{13}$$

where Z_{c} is the fixed CoM height.

3. Model implementation

The model in (11) has been developed exploiting the characteristics of the system and adopting some approximation based on the three assumptions reported in Section 2.3. These approximations allow to reduce the complexity of the equations but at the same time they introduce an error in the model. The model equations can be rearranged as follows:

$$\xi_1 = \ddot{x} = -\alpha_1 \frac{K_{xx}}{m}(x - x_{\mathrm{r}}) - \beta_1 \frac{K_{xz}}{m}(z - z_{\mathrm{r}}) - \gamma_1 \frac{D_{xx}}{m}\dot{x} - \varepsilon_1 \frac{D_{xz}}{m}\dot{z}, \tag{14}$$

$$\xi_2 = \ddot{z} = -g - \alpha_2 \frac{K_{zx}}{m}(x - x_{\mathrm{r}}) - \beta_2 \frac{K_{zz}}{m}(z - z_{\mathrm{r}}) - \gamma_2 \frac{D_{zx}}{m}\dot{x} - \varepsilon_2 \frac{D_{zz}}{m}\dot{z}, \tag{15}$$

where α_i, β_i, γ_i and ε_i, i={1, 2} are parameters inserted to compensate model errors due to the approximation used as well as other errors from the identification of the joint stiffness and damping parameters.

Equations (14) and (15) describe the robot behavior during single support. In this phase, the robot stands on one leg therefore the robot movement is mostly affected from the compliant joints of the supporting leg. Otherwise, during double support phase, both feet are on the ground hence the compliant joints of both legs contribute to the robot movement. To take into account the effect of the second leg during this phase, the same procedure used to derive equations (9) can be reiterated for the other leg and included in the model equation. In order to reduce the complexity of the model, another approach has been adopted. Assuming that during the double support phase the two feet on the ground do not move relative to each other, the forces developed from the two legs are different because of the different configuration of the legs and different displacements. The effect of the second leg has been included to the model scaling equations (14) and (15) as follows:

$$\xi_1 = \ddot{x} = \eta_1 \left(-\alpha_1 \frac{K_{xx}}{m}(x - x_{\mathrm{r}}) - \gamma_1 \frac{D_{xx}}{m}\dot{x} - \beta_1 \frac{K_{xz}}{m}(z - z_{\mathrm{r}}) - \varepsilon_1 \frac{D_{xz}}{m}\dot{z}\right), \tag{16}$$

$$\xi_2 = \ddot{z} = -g + \eta_2 \left(-\alpha_2 \frac{K_{zx}}{m}(x - x_{\mathrm{r}}) - \gamma_2 \frac{D_{zx}}{m}\dot{x} - \beta_2 \frac{K_{zz}}{m}(z - z_{\mathrm{r}}) - \varepsilon_2 \frac{D_{zz}}{m}\dot{z} \right). \tag{17}$$

Equations (16) and (17) are used during the double support phase. The scaling coefficients η_1, η_2 are computed from experimental data which evaluates the *x* and *z* forces measured by the force/torque sensors mounted at the feet of the robot [14].

4. Optimization of the model using reinforcement learning

The proposed compact model captures well many of the characteristics of the passively compliant robot. However, it does not perfectly predict the behaviour of the robot, which is due to two main reasons: (i) the simplifying assumptions, which are necessary to keep the model compact, but also introduce approximation errors, and (ii) measurement inaccuracies when estimating the physical properties of the robot, e.g. the stiffness of the springs, etc. A common simplification assumption is, for example, that the left and the right leg have the same properties and therefore behave identically. In reality, this is not entirely true, because of the complexity of series elastic actuation, where it is normal to observe differences between the right and left leg's motors, passive compliance unit, friction, and so on.

One way to minimize the modelling error is to make the model more complex. This, however, would diminish the advantage of having a compact model, and would impede its use as a fast predictor of the robot's behaviour instead of the robot itself. Therefore, in this paper we concentrate our effort on optimizing the proposed compact model in order to achieve the best modelling precision without unnecessarily increasing the model complexity.

More concretely, by optimization of the model we mean the search for optimal values of some important parameters of the model. From equations (14)-(17) we have identified 8 important parameters whose values are crucial for the performance of the model. These parameters are as follows: α_1 and γ_1 affect the relationship between the force generate along x direction and the position and velocity of the CoM in the same direction, β_2 and ε_2 affect the relationship between the force generate along z direction and the position and velocity of the CoM in the same direction, finally β_1, ε_1, α_2 and γ_2 affect the coupling between the orizontal and vertical movement of the CoM.

Initially, the parameter values are all set to be equal to 1.0. The goal of the optimization is to find other values for these parameters which reduce the overall model error. The model error is estimated using a ground truth data set, recorded from real-world experiments with COMAN, where the reference trajectories are known, and the actual response trajectories are compared to the model output. Using this data set, we have defined a cost function to be minimized, equal to the mean squared error of the model prediction with respect to the actual response trajectories.

Many alternative optimization approaches exist, which can be used to minimize this cost function. In this paper we have selected a reinforcement learning

approach, based on direct policy search. In this approach, the policy is parameterized by a set of parameters, and the reinforcement learning algorithm is trying to optimize their values by performing a sequence of trials and evaluating its performance using the defined cost function (which is a form of reward function).

In particular, we have selected the POWER algorithm (POlicy learning by Weighting Exploration with the Returns [18]) for implementing the optimization, for the following reasons: (i) it does not require a learning rate parameter; (ii) it re-uses efficiently trials using importance sampling; (iii) it has been already applied successfully on the COMAN robot for another task, to optimize the walking gait by varying the center-of-mass height and thus reduce the energy consumption [19].

The optimization algorithm has been executed on reference trajectories with duration of 30 s (30 000 samples) and the corresponding trajectories performed by the robot computed through forward kinematics. The gait parameters are sl=0.03, T_{ss}=0.5 s and T_{db}= 0.2 s.

Table 1. Parameter values after the optimization

Parameter	α_1	β_1	γ_1	ε_1	α_2	β_2	γ_2	ε_2
Value	0.9827	1.0357	1.0560	1.1335	0.9792	0.9561	1.0072	1.1053

Table 1 contains the optimal values of the parameters found by the algorithm. The results from the optimization show that the modelling error can indeed be reduced by only changing the values of the selected eight parameters ($\alpha_1, \beta_1, \gamma_1, \varepsilon_1, \alpha_2, \beta_2, \gamma_2, \varepsilon_2$). In Fig. 4 the results of the model before and after the optimization process are compared.

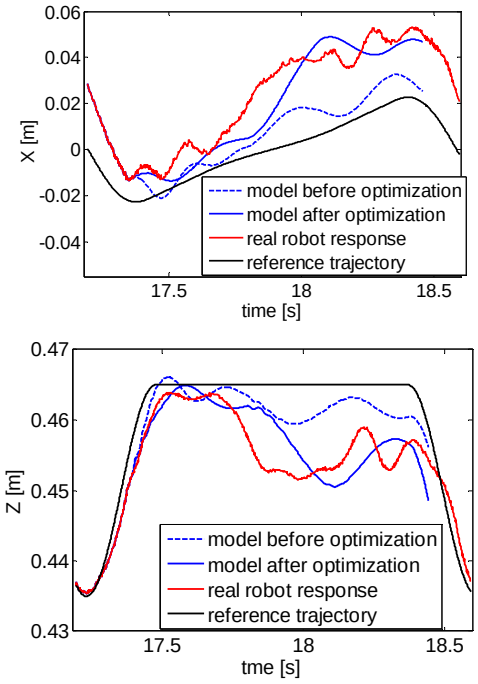

Fig. 4. Comparison of the model results before and after optimization

The base assumptions of the model reduce the precision of the inertia representation more than the gravity. The movement along the walking direction is mostly affected by the inertia of the system because there are important accelerations and decelerations. Instead vertical movement is mostly affected by gravity force. The consequence is that the movement along x directions benefit of the optimization procedure more than z direction.

5. Conclusion

In this work a reduced model of the dynamic of the CoM motion of the robot has been described. The model is based on some assumptions that reduce the complexity of the equations but at the same reduce the precision of the results. To improve the performances of the model some parameters have been inserted in the model equations in order to compensate the errors due to the reductions. The value of these parameters has been optimized using learning technique. An improvement of the model performance has been reached.

The set of parameters found has been also used to perform walking with different gaits than the one used in the optimization process. Also, in these cases the performance reached by the optimized model is better than the original model.

Acknowledgment. This work is supported by the FP7 European project AMARSI (ICT-248311).

References

1. Hirose, M., Y. Haikawa, T. Takenaka, K. Hirai. Development of Humanoid Robot ASIMO. – In: Proc. IEEE/RSJ IROS'2001, Workshop2.
2. Hirai, K., M. Hirose, Y. Haikawa, T. Takenaka. The Development of Honda Humanoid Robot. – In: Proc. of IEEE ICRA'1998, 1321-1326.
3. Akachi, K., K. Kaneko, N Kanehira, S. Ota, G. Miyamori, M. Hirata, S. Kajita, F. Kanehiro. Development of Humanoid Robot HRP-3P. – In: Proc. of IEEE-RAS Int. Conf. on Humanoid Robots, 50-55.
4. Ogura, Y., H. Aikawa, K. Shimomura, A. Morishima, H. Lim, A. Takanishi. Development of a New Humanoid Robot WABIAN-2. – In: Proc. of IEEE ICRA'2006, 76-81.
5. Tsagarakis, N. G., G. Metta, G. Sandini, D. Vernon, R. Beira, F. Becchi, L. Righetti, J. S. Victor, A. J. Ijspeert, M. C. Carrozza, D. G. Caldwell. iCub: The Design and Realization of an Open Humanoid Platform for Cognitive and Neuroscience Research. – Advanced Robotics, Vol. **21**, 2007, No 10, 1151-1175.
6. Tsagarakis, N. G., B. Vanderborght, M. Laffranchi, D. G. Caldwell. The Mechanical Design of the New Lower Body for the Child Humanoid Robot "iCub". – In: IEEE/RSJ International Conference on Intelligent Robots and Systems, 2009, 4962-4968.
7. Tsagarakis, N. G., F. Becchi, M. Singlair, G. Metta, D. G. Caldwell, G. Sandini. Lower Body Realization of the Baby Humanoid-"iCub". – In: IEEE/RSJ International Conference on Intelligent Robots and Systems, 2007, 3616-3622.
8. Lohmeier, S., T. Buschmann, H. Ulbrich, F. Pfeiffer. Modular Joint Design for Performance Enhanced Humanoid Robot LOLA. – In: Proc. of IEEE ICRA'2006, 88-93.
9. Yamaguchi, J., E. Soga, S. Inoue, A. Takanishi. Development of a Bipedal Humanoid Robot – Control Method of Whole Body Cooperative Dynamic Biped Walking. – In: Proc. of IEEE ICRA'1999, 368-374.

10. P a r k, I. W., J. Y. K i m, J. L e e, J. H. O h. Mechanical Design of the Humanoid Robot Platform HUBO. – Journal of Advanced Robotics, Vol. **21**, 2007, No 11, 1305-1322.
11. T s a g a r a k i s, N. G., M. S i n g l a i r, F. B e c c h i, G. M e t t a, G. S a n d i n i, D. C a l d w e l l. Lower Body Design of the "iCub" a Human-Baby Like Crawling Robot. – IEEE Humanoids, 2006, 450-455.
12. T s a g a r a k i s, N., Z. L i, J. S a g l i a, D. G. C a l d w e l l. The Design of the Lower Body of the Compliant Humanoid Robot "cCub". – In: IEEE ICRA'2011.
13. T s a g a r a k i s, N., M. L a f f r a n c h i, B. V a n d e r b o r g h t, D. C a l d w e l l. A Compact Soft Actuator Unit for Small Scale Human Friendly Robots. – In: IEEE ICRA'2009, 4356-4362.
14. S p o n g, M. Modeling and Control of Elastic Joint Robots. – Trans. ASME : J. Dyn. Syst., Meas., Control, Vol. **109**, 1987, 310-319.
15. C o l a s a n t o, L., N. G. T s a g a r a k i s, D. G. C a l d w e l l. A Compact Model for the Compliant Humanoid Robot COMAN. – In: IEEE International Conference on Biomedical Robotics and Biomechatronics, Rome, Italy, 2012.
16. C h e n, S., J. K a o. Simulation of Conservative Congruence Transformation Conservative Properties in Joint and Cartesian Spaces. – In: IEEE International Conference on Intelligent Robotics and Automation, 2001, 3356-3363.
17. K a j i t a, S., M. M o r i s a w a, K. M i u r a, S. N a k a o k a, K. H a r a d a, K. K a n e k o, F. K a n e h i r o, K. Y o k o i. Biped Walking Stabilization Based on Linear Inverted Pendulum Tracking. – In: IEEE/RSJ International Conference on Intelligent Robots and Systems, Taipei, Taiwan, 2010, 4489-4496.
18. K a j i t a, S., F. K a n e h i r o, K. K a n e k o, K. F u j i w a r a, K. H a r a d a H. H i r u k a v a, K. Y o k o i. Biped Walking Pattern Generation by using Preview Control of Zero-Moment Point. – In: IEEE/RSJ International Conference on Intelligent Robots and Systems, Las Vegas, USA, 2003, 1620-1626.
19. K o b e r, J., J. P e t e r s. Learning Motor Primitives for Robotics. – In: Proc. IEEE Intl. Conf. on Robotics and Automation (ICRA), May 2009, 2112-2118.
20. K o r m u s h e v, P., B. U g u r l u, S. C a l i n o n, N. G. T s a g a r a k i s, D. G. C a l d w e l l. Bipedal Walking Energy Minimization by Reinforcement Learning With Evolving Policy Parameterization. – In: Proc. IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS), San Francisco, USA, September 2011, 318-324.