Vision-Based Trajectory Control for Humanoid Navigation

Giuseppe Oriolo, Antonio Paolillo, Lorenzo Rosa, Marilena Vendittelli

Abstract— We address the problem of robustly tracking a desired workspace trajectory with a humanoid robot. The proposed solution is based on the suitable definition of a controlled output, which represents an averaged motion of the torso after cancellation of the sway oscillation. In particular, two different techniques are presented for extracting the averaged motion. For control design purposes, a unicycle-like model is associated to the evolution of this output. The feedback loop is then closed using a vision-based odometric localization method to estimate the torso motion. The proposed approach is validated through comparative experiments on the humanoid robot NAO.

I. Introduction

In mobile robot control, feedback tracking of trajectories that have been planned to satisfy certain objectives (such as obstacle avoidance, minimization of time and traveled distance, or mechanical balance) guarantees that these are robustly achieved during motion execution, in spite of the presence of uncertainties and perturbations that invariably affect actual operation. It is well known (see for example the experiments reported in [1]) that open-loop execution of commanded motion can be dramatically inaccurate in the case of humanoid robots, due to their complex kinematic structure that amplifies mechanical imprecisions, to non-ideal interaction with the ground (e.g., foot slippage), and so on.

The trajectory tracking problem for humanoids in itself has received very little attention in the literature. Reacting to external perturbations and environment changes are related problems usually addressed by replanning [2], [3], [4], [5] or on-line adapting the feet position [6], [7], [8]. In particular, a problem which is conceptually close to the trajectory tracking considered here is dealt with in [9], where the planned footsteps sequence is adapted so as to compensate for the effect of perturbations on the center of mass trajectory at execution time.

In this paper, we consider the problem of tracking a generic workspace trajectory with a humanoid robot. The main difficulty in this case is to define an appropriate output for tracking. In fact, no fixed point on the robot is eligible for this purpose, in view of the sway oscillation due to the walking gait which is always superimposed to the motion of the robot. This natural oscillation would be interpreted as an error and therefore corrected by any trajectory tracking controller, with obvious negative effects on the actually generated motion.

The authors are with the Dipartimento di Ingegneria Informatica, Automatica e Gestionale, Sapienza Università di Roma, via Ariosto 25, 00185 Roma, Italy. Email: {oriolo, paolillo, rosa, vendittelli}@dis.uniroma1.it

To solve this problem, we propose to use as controlled output a virtual reference point that represents the averaged motion of the torso after the sway oscillation has been canceled. Accordingly, we solve the tracking control problem for this output, delegating the generation of footsteps to the humanoid built-in walking engine which we assumed to be available. This is a convenient approach for execution of navigation tasks where only the Cartesian motion of the robot is of interest. Within this formulation of the trajectory tracking problem, a unicycle-like mobility model is associated to the virtual reference point for control design purposes. The vision-based odometric localization method developed in [1] is used to estimate the torso motion and close the feedback loop.

The paper is organized as follows. In the next section we provide a general description of the proposed approach. Section III summarizes the main features of the odometric localization method proposed in [1]. Two different averaging techniques are then proposed for sway motion cancellation in Section IV, while the adopted mobility model with the associated tracking control are described in Sect. V. Section VI presents comparative control experiments on the NAO humanoid robot tracking a linear and a sigmoid trajectory. Section VII concludes the paper.

II. Proposed approach

The basic idea of this paper is to track the given workspace trajectory with a 'virtual' reference point representing the averaged motion of the humanoid torso. This averaged torso motion is obtained after cancellation of the *sway motion*, i.e. the transversal oscillation that the torso naturally performs during locomotion.

For the virtual reference point, we will adopt the mobility model of a unicycle, which has been established as an appropriate model for long-distance human walking [10], [11]. As in [12], we shall use this mobility model for control design and choose driving and steering velocity inputs so as to achieve our navigation objective. We will assume that the humanoid is equipped with a walking engine that accepts these velocities as high-level commands and converts them to an appropriate footstep plan. This is the case, for example, of the NAO robot which will be used for our experiments.

The block diagram of Fig. 1 illustrates the proposed approach. In particular, the relevant signals are:

- $\boldsymbol{w} = (x_t, y_t, \theta_t)$, where (x_t, y_t) are the coordinates of the torso on the ground plane and θ_t is its yaw angle;
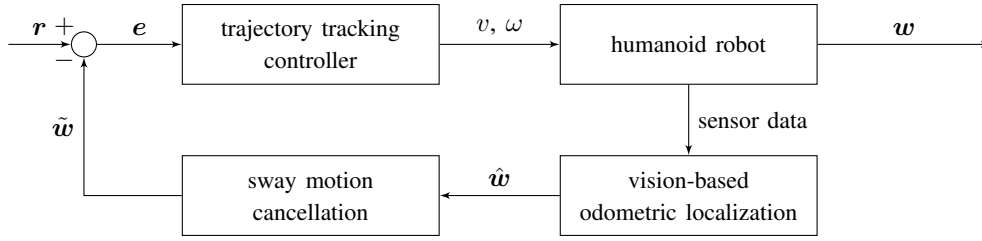
Fig. 1.   High-level block diagram of the proposed control approach.

- $\hat{\boldsymbol{w}} = (\hat{x}_t, \hat{y}_t, \hat{\theta}_t)$, the estimate of $\boldsymbol{w}$ generated by a vision-based odometric localization algorithm;
- $\tilde{\boldsymbol{w}} = (x, y, \theta)$, the averaged version of $\hat{\boldsymbol{w}}$ in which the effect of the sway motion has been canceled;
- $\boldsymbol{r} = (x_d, y_d, \theta_d)$, the assigned reference trajectory to be tracked with $\tilde{\boldsymbol{w}} = (x, y, \theta)$.

Note that the torso roll and pitch angles are not considered in this high-level control loop.

For odometric localization, we shall adopt the technique introduced in [1]. Sway motion cancellation can be achieved in different ways and is a specific contribution of this paper. Finally, any trajectory tracking controller devised for unicycle-like robots is eligible for the implementation of the controller block.

In the following sections, we describe in some detail the individual blocks of the control scheme.

III. Vision-based Odometric Localization

Odometric localization of humanoid robots is a challenging problem due to complex kinematics of these robots and the hybrid nature of locomotion. In [1] we have proposed a vision-based odometric localization method that provides an estimate of the torso pose (position and orientation). Such estimate is produced by an Extended Kalman Filter, according to the conceptual scheme shown in Fig. 2. The state of the filter is therefore the pose of the torso, while the measured variables are the pose of the head and the orientation of the torso.

The prediction of the torso pose is obtained through the robot forward kinematics from the support foot, using data from the encoders of the support leg joints. The head pose is predicted using the kinematic map from the torso to the head and the neck encoders data. This prediction is then corrected using the measurement of the head pose provided by the PTAM algorithm [13], which acts here as a sensor, and the torso orientation measured by the onboard Inertial Measurement Unit (IMU). An asynchronous update mechanism triggered by the pressure sensors allows switching to the current support foot for the correct kinematic computations.

Note that, while the filter designed in [1] provides the whole pose of the torso, for the trajectory tracking problem we use only part of the estimate, namely the coordinates of the projection on the ground of the torso position and the torso yaw angle.

In [1] the localization system was tested on a batch of data collected during locomotion. To implement the control

scheme of Fig. 1, we had to realize an on-line implementation of the filter; this required the solution of some technical issues mainly related to the synchronization of the information provided by the various sensors.

IV. Sway Motion Cancellation

Two different techniques have been developed to achieve sway motion cancellation and isolate an averaged motion of the humanoid. The first proceeds from the observation that the swaying oscillation is a relatively high-frequency phenomenon, and therefore it may be removed by a suitable low-pass filter. The second technique uses instead a geometric projection procedure to cancel the lateral movement of the torso during locomotion.

A. Low-pass filtering

To perform sway motion cancellation via low-pass filtering, one needs a preliminary analysis of the walking motion in the frequency domain. The experimental platform which we have used to collect data is the humanoid robot NAO by Aldebaran Robotics. In particular, we have moved the robot along a path similar to that of Figs. 5–6 and used MATLAB's function FFT to compute the Fourier transform of the torso position coordinates $\hat{x}_t$ and $\hat{y}_t$ as estimated by the odometric localization system described in Sect. III. The resulting magnitude plots, shown in Fig. 3, clearly indicates that the frequency associated to the sway motion is very close to 1 Hz. As a consequence, a first-order continuos time filter was used with a cut-off frequency of 0.8 Hz.

While the first two components x, y of the averaged vector $\tilde{\boldsymbol{w}}$ are directly the low-pass filtered versions of $\hat{x}_t$ and $\hat{y}_t$, its last component, which represents the averaged orientation, is computed as the tangent to the averaged x-y trajectory, i.e.

$$\theta = \text{Atan2}(\dot{y}, \dot{x}).$$

The rationale of this definition of the averaged orientation is consistency: since we intend to use a unicycle controller for trajectory tracking, it appears reasonable to guarantee that $\tilde{w}$ evolves as the configuration of a unicycle, and in particular that θ obeys the pure rolling constraint. In practice, in the above formula $\dot{x}$ and $\dot{y}$ are approximately computed from the filter output via backward difference.

A possible limitation of the above approach for sway motion cancellation is the non-selectivity of the filtering action: all harmonic components above the cut-off frequency
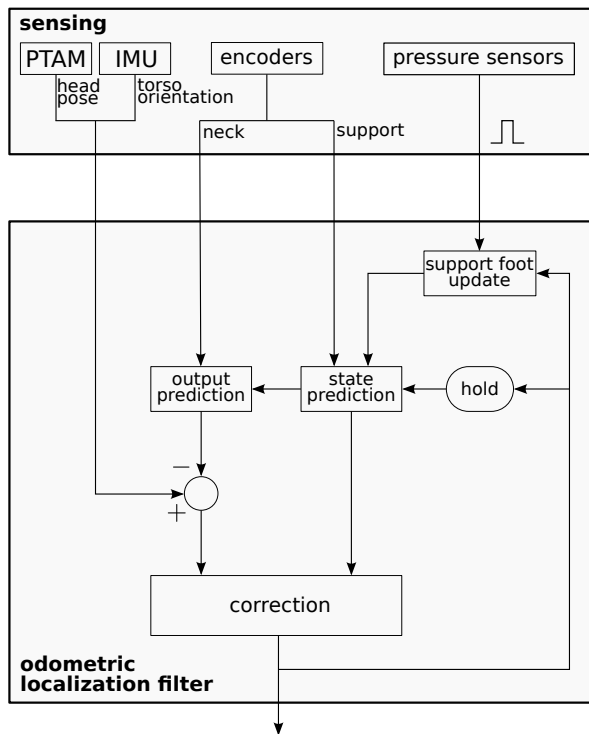
Fig. 2. Block scheme of the vision-based odometric localization system. The EKF state is the full pose of the torso. Only part of this estimate is used for trajectory tracking.
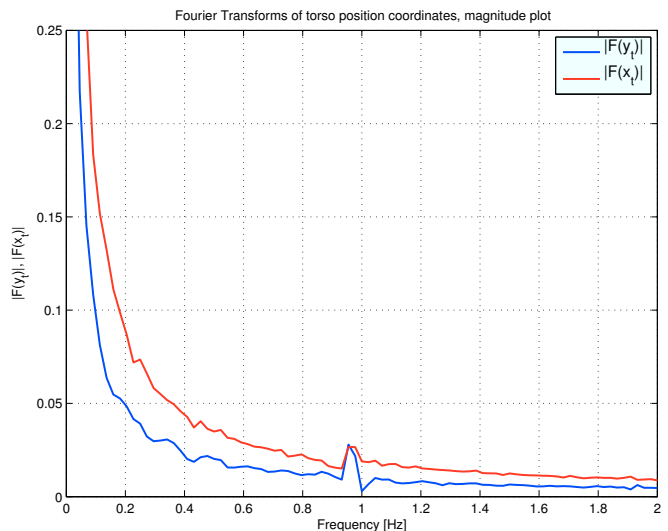

Fig. 3. Magnitude plots of the Fourier trasforms of the estimated torso coordinates $\hat{x}_t$ and $\hat{y}_t$ along a typical walking trajectory.

will be removed together with the sway motion, and in principle this may result in a 'smoothing' of the reconstructed averaged motion with respect to the actual one. In addition, one may expect the low-pass filter to introduce a small delay between the torso estimates and the averaged motion.

B. Geometric projection

The second technique to cancel the sway motion relies on a geometric projection procedure. The basic idea is to compute the lateral displacement of the torso with respect to the current support foot and to use this information to compensate the sway motion associated to walking.

With reference to Fig. 4, consider the following construction. From the origin of the frame attached to the support foot, move in the inner direction orthogonal to the foot of a distance $\delta_{\mathcal{P}}$, the half distance between the robot feet in the standing rest position; call S the reached point. Now consider a vertical plane $\mathcal{P}$ that passes through S and is oriented as the torso; i.e., the intersection of $\mathcal{P}$ with the ground plane is a line whose orientation in the world frame is θ_t. Note that $\mathcal{P}$ represents the robot sagittal plane. For NAO robot, $\delta_{\mathcal{P}}$ is equal to 0.05 m.

The first two components x, y of the averaged vector $\tilde{\boldsymbol{w}}$ are obtained as the first two coordinates of the orthogonal projection of the origin of the torso frame, whose position is (x_t, y_t), on the sagittal plane $\mathcal{P}$; whereas its last component θ, which represents the averaged orientation, is obtained from x, y as in the previous technique.

In principle, this second approach for sway motion cancellation should be more expensive than the first from a computational viewpoint, due to the involved geometric calculations. However, all the quantities needed to implement the above construction are already derived and available in the odometric localization module; therefore, the additional cost for our integrated scheme of Fig. 1 is actually negligible.

C. Experimental comparison

Figures 5 and 6 show the application of the two above techniques for cancelling sway motion during a typical walk. Both are clearly successful and extract rather well the averaged motion. A closer look at the plots reveals that with the first technique the residual oscillation is slightly larger than with the second, suggesting a further reduction of the cut-off frequency of the filter; also, a small delay is present — as expected — between the two signals. With the second technique, the residual oscillation is very limited and no delay is present.

V. TRAJECTORY CONTROL

Both the sway motion cancellation techniques of the previous section provide an averaged motion in the form of a vector $\tilde{\boldsymbol{w}} = (x, y, \theta)$, where (x, y) represent the Cartesian coordinates of a virtual reference point for the humanoid and θ is related to the tangent to the trajectory of this point. As explained in Sect. II, we assume a unicycle-like mobility model for the virtual reference point, i.e.,

$$\begin{aligned} \dot{x} &= v\cos\theta \\ \dot{y} &= v\sin\theta \\ \dot{\theta} &= \omega, \end{aligned}$$

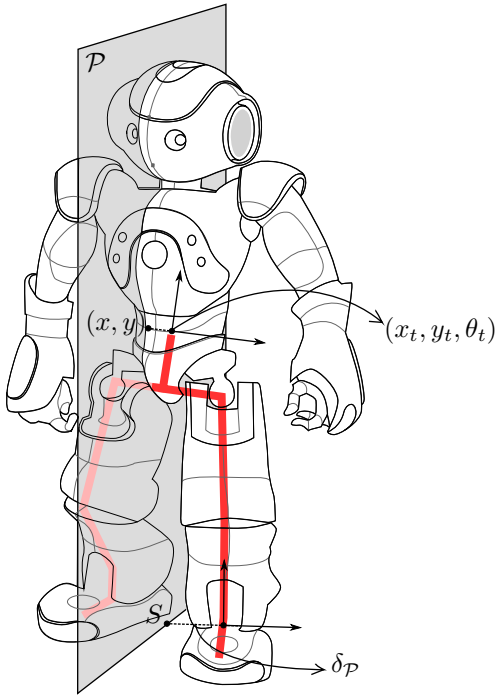where v and ω are, respectively, the steering and driving velocities sent to the locomotion controller of the humanoid.

Fig. 4.   Geometric representation of the projection procedure.

Define the current error w.r.t. the desired trajectory as

$$\boldsymbol{e} = \begin{pmatrix} e_x \\ e_y \\ e_\theta \end{pmatrix} = \begin{pmatrix} x_d - x \\ y_d - y \\ \theta_d - \theta \end{pmatrix},$$

where $(x_d, y_d, \theta_d) = \boldsymbol{r}$ is the reference signal associated to the desired trajectory in the Cartesian workspace.

At this point, any trajectory tracking controller for unicycle-like robots can be used to close the loop. In particular, in our implementation we have used the scheme proposed in [14], which consists of a nonlinear time-invariant control law:

$$v = v_d \cos e_\theta + k_1(e_x \cos\theta + e_y \sin\theta)$$
$$\omega = \omega_d + k_2 v_d \frac{\sin e_\theta}{e_\theta}(e_y \cos\theta - e_x \sin\theta) + k_3 e_\theta.$$

Here, the control gains are chosen as follows:

$$k_1 = k_3 = 2a\sqrt{\omega_d^2(t) + bv_d^2(t)}, \quad k_2 = b,$$

with $a \in (0,1)$ and $b > 0$, whereas the feedforward commands are computed from the desired trajectory:

$$v_d(t) = \sqrt{\dot{x}_d^2(t) + \dot{y}_d^2(t)}$$
$$\omega_d(t) = \frac{\ddot{y}_d(t)\dot{x}_d(t) - \ddot{x}_d(t)\dot{y}_d(t)}{\dot{x}_d^2(t) + \dot{y}_d^2(t)}.$$

VI. Experimental Results

As already mentioned, the experimental platform we have used to validate the proposed control scheme is the humanoid robot NAO by Aldebaran Robotics.
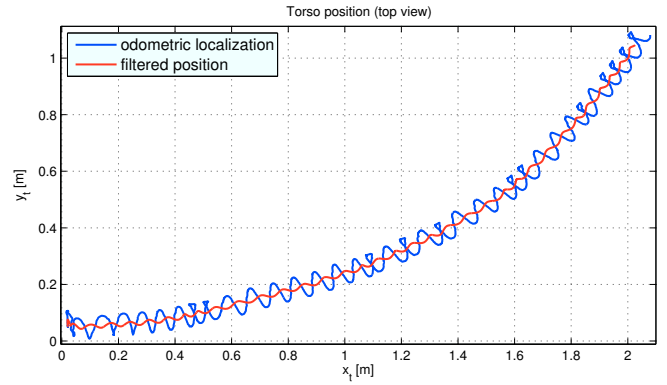

Fig. 5.   Sway motion cancellation via low-pass filtering: Odometric localization estimate of the torso position vs. its filtered version.
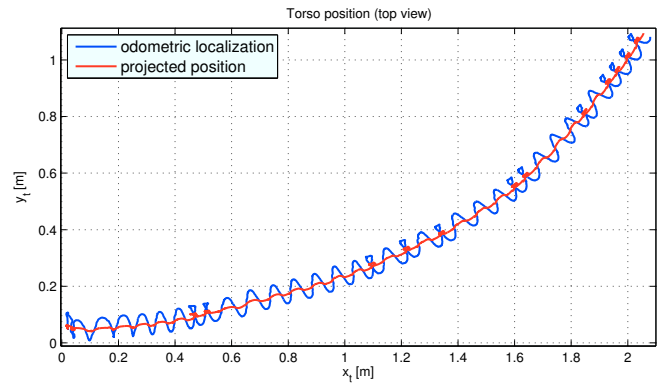

Fig. 6.   Sway motion cancellation via geometric projection: Odometric localization estimate of the torso position vs. its projected version.

The NAO sensors used by the odometric localization system are: the CMOS digital camera with a $72.6°$ field of view mounted on the forehead; the IMU located in the chest, which provides a measure of the torso roll and pitch angles; the magnetic rotary encoders with $0.1°$ resolution available at the joints; and the force sensitive resistors located under the feet, that measure the applied pressure.

Since odometric localization runs at a nominal rate of 100 Hz, this is also the control rate at which the driving and steering velocity inputs are updated. These commands are then sent to the robot using the NAO APIs, and in particular the built-in *move* function. Since the most recent command overrides all previous commands, this function can be called with arbitrary rate, thus providing a convenient mechanism for real-time implementation of a high-level control loop.

The experiments are performed in an indoor environment. No special structuring was needed for localization, as the PTAM component of the EKF algorithm successfully uses the natural features of the environment. The metric scale is recovered from kinematic computations by performing a preliminary motion on the spot before starting the navigation task, as explained in [1]. To maximize the performance of the localization algorithm, the computations are performed on-line on a remote PC, although this choice introduces some

delay due to data transmission.

Video clips of the presented experiments are shown in the multimedia attachment to this paper.

A. Linear trajectory

In the first experiment, the desired trajectory is a line starting from the robot initial position. The desired speed is $v_d = 0.5$ m/s, while the control parameters were set to $a = 0.8$ and $b = 10.0$. For the sake of comparison, we present results obtained with both sway motion cancellation techniques.

Figures 7–8 report the results obtained using low-pass filtering for sway motion cancellation. In particular, the first shows the actual trajectory of the torso, as estimated by our odometric localization algorithm (no ground truth is available), while the second shows the controlled variable (the filtered position); in both figures, the desired trajectory is also shown as reference. A transient error is associated to the starting phase of the walking motion. As soon as the locomotion becomes regular, the controller is more effective and the error quickly decays. The robot is then able to closely track the desired trajectory.

Figures 9–10 show the corresponding results obtained using geometric projection for sway motion cancellation. The performance is similar to the previous case, although low-pass filtering results in a slightly better behavior. For the latter, a stroboscopic motion representation is shown in Fig. 11.

To allow a quantitative comparison, we computed the root mean square e_{rms} of the cartesian error along the generated motions, obtaining $e_{\mathrm{rms}} = 0.0330$ m for the low-pass filtering and $e_{\mathrm{rms}} = 0.0808$ m for the geometric projection.

B. Sigmoid trajectory

The sigmoid trajectory is composed by two straight line segments, to be traced with a desired linear speed of $v_d = 0.05$ m/s, connected by two arcs of circle of radius 0.25 m, to be traced with a desired angular speed of 0.10 rad/s. The values of the control parameters were set at $a = 0.9$ and $b = 100.0$ in this case.

As shown in Figs. 12–15, the results are satisfactory for both sway cancellation methods, again with a slight advantage for low-pass filtering that achieves a slightly smoother motion. For the latter, a stroboscopic motion representation is shown in Fig. 16.

In this case we obtained $e_{\mathrm{rms}} = 0.0186$ m for the low-pass filtering and $e_{\mathrm{rms}} = 0.0191$ m for the geometric projection.

VII. CONCLUSIONS

We have presented a feedback control scheme for vision-based trajectory tracking with humanoids. A vision-based odometric localization system provides an estimate of the torso pose which is used as feedback signal to close the loop. Two different methods are devised for sway motion cancellation. Experiments performed on the humanoid robot NAO have been presented to validate our approach.
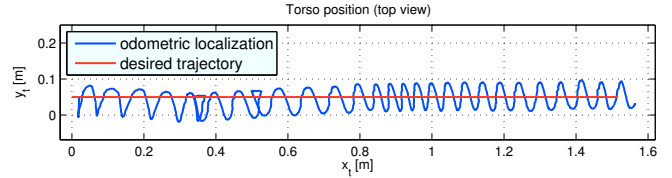

Fig. 7. Linear trajectory tracking using low-pass filtering for sway motion cancellation: trajectory of the torso vs. desired humanoid trajectory.
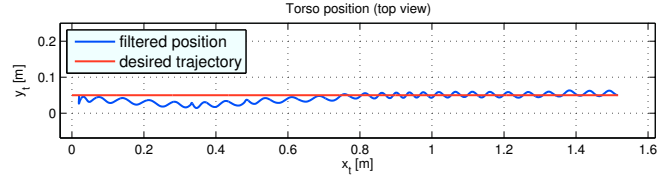

Fig. 8. Linear trajectory tracking using low-pass filtering for sway motion cancellation: the controlled variable (the filtered position) vs. the reference signal.
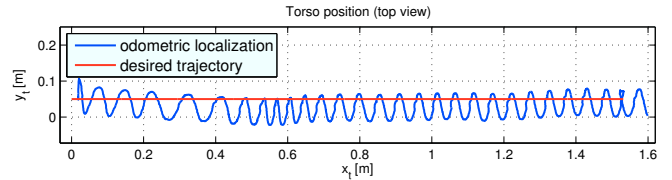

Fig. 9. Linear trajectory tracking using geometric projection for sway motion cancellation: trajectory of the torso vs. desired humanoid trajectory.
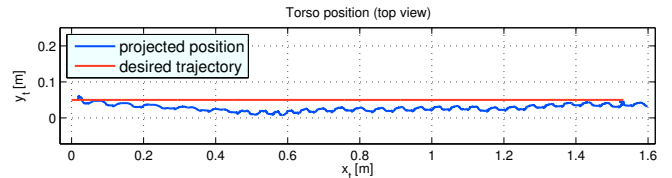

Fig. 10. Linear trajectory tracking using geometric projection for sway motion cancellation: the controlled variable (the projected position) vs. the reference signal.
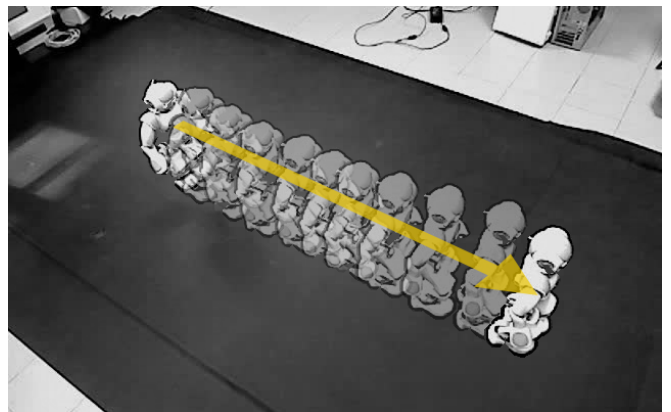

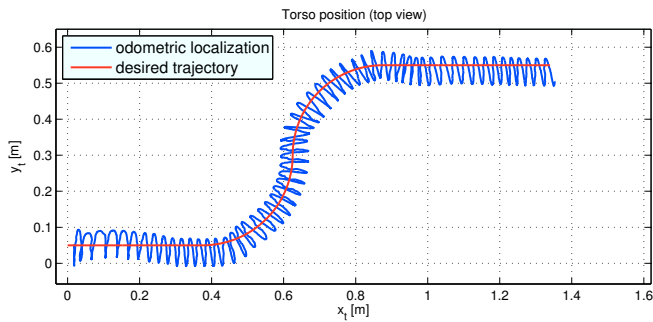Fig. 11. Line trajectory tracking: Stroboscopic motion

Fig. 12. Sigmoid trajectory tracking using low-pass filtering for sway motion cancellation: trajectory of the torso vs. desired humanoid trajectory.
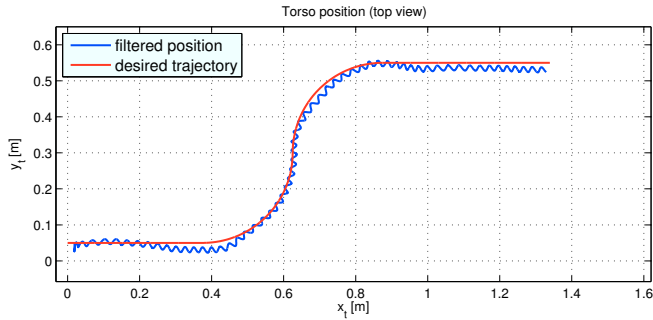

Fig. 13. Sigmoid trajectory tracking using low-pass filtering for sway motion cancellation: the controlled variable (the filtered position) vs. the reference signal.
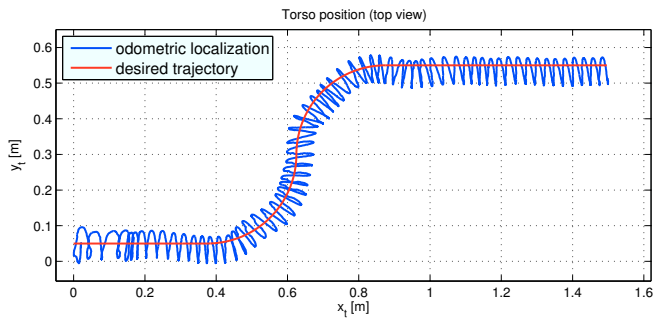

Fig. 14. Sigmoid trajectory tracking using geometric projection for sway motion cancellation: trajectory of the torso vs. desired humanoid trajectory.
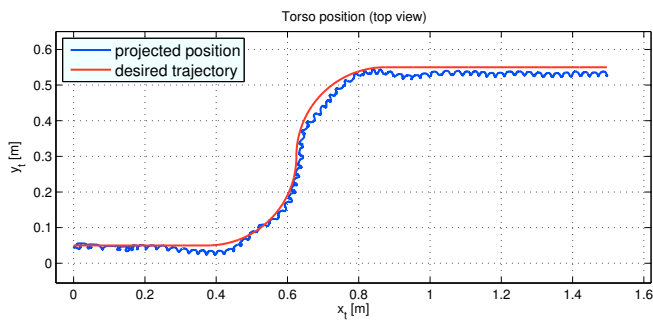

Fig. 15. Sigmoid trajectory tracking using geometric projection for sway motion cancellation: the controlled variable (the projected position) vs. the reference signal.
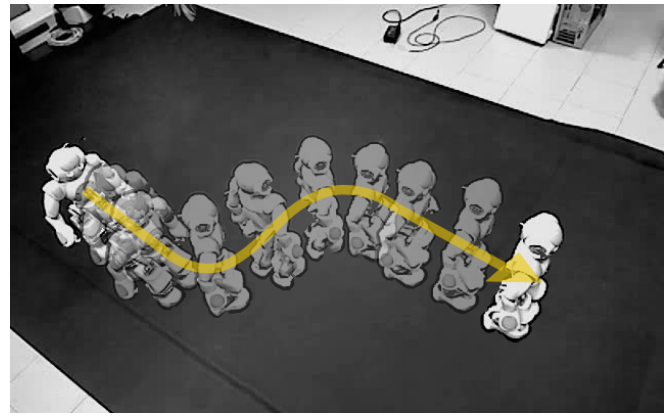

Fig. 16. Sigmoid trajectory tracking: Stroboscopic motion.

REFERENCES

[1] G. Oriolo, A. Paolillo, L. Rosa, and M. Vendittelli, "Vision-based odometric localization for humanoids using a kinematic EKF," in *IEEE-RAS International Conference on Humanoid Robots*, pp. 152–158, 2012.

[2] P. Michel, J. Chestnutt, J. Kuffner, and T. Kanade, "Vision-guided humanoid footstep planning for dynamic environments," in *5th IEEE-RAS International Conference on Humanoid Robots*, p. 1318, 2005.

[3] P. Michel, J. Chestnut, S. Kagami, K. Nishiwaki, J. Kuffner, and T. Kanade, "Online environment reconstruction for biped navigation," in *2006 IEEE International Conference on Robotics and Automation*, p. 3089309, 2006.

[4] J. Chestnut, "Navigation and gait planning," in *Motion planning for humanoid robots* (K. Harada, E. Yoshida, and K. Yokoi, eds.), pp. 1–28, 2019.

[5] L. Baudouin, P. Perrin, T. Moulard, O. Stasse, F. Lamiraux, and E. Yoshida, "Real-time replanning using 3d environment for humanoid robot," in *11th IEEE-RAS International Conference on Humanoid Robots*, pp. 584–589, 2011.

[6] K. Nishiwaki and S. Kagami, "High frequency walking pattern generation based on preview control of zmp," in *2006 IEEE International Conference on Robotics and Automation*, p. 26672672, 2006.

[7] A. Herdt, H. Diedam, P.-B. Wieber, D. Dimitrov, K. Mombaur, and M. Diehln, "Online walking motion generation with automatic foot step placements," *Advanced Robotics*, vol. 24, no. 5-6, p. 719737, 2010.

[8] D. Dimitrov, A. Paolillo, and P.-B. Wieber, "Walking motion generation with online foot position adaptation based on l1- and l1-norm penalty formulations," in *2011 IEEE International Conference on Robotics and Automation*, pp. 3523–3529, 2011.

[9] T. Moulard, F. Lamiraux, and O. Stasse, "Trajectory following for legged robots," in *International Conference on Biomedical Robotics and Biomechatronics*, 2012.

[10] T.-V.-A. Truong, D. Flavigne, J. Pettre, K. Mombaur, and J.-P. Laumond, "Reactive synthesizing of human locomotion combining nonholonomic and holonomic behaviors," in *3rd IEEE/RAS-EMBS International Conference on Biomedical Robotics and Biomechatronics*, pp. 632–637, 2010.

[11] K. Mombaur, A. Truong, and J.-P. Laumond, "From human to humanoid locomotion – an inverse optimal control approach," *Autonomous Robots*, vol. 28, pp. 369–383, 2010.

[12] A. Faragasso, G. Oriolo, A. Paolillo, and M. Vendittelli, "Vision-based corridor navigation for humanoid robots," in *2013 IEEE International Conference on Robotics and Automation*, pp. 152–158, 2013.

[13] G. Klein and D. Murray, "Parallel tracking and mapping for small AR workspaces," in *6th IEEE and ACM Int. Symp. on Mixed and Augmented Reality*, pp. 225–234, 2007.

[14] C. Samson, "Time-varying feedback stabilization of car-like wheeled mobile robots," *International Journal of Robotics Research*, vol. 12, no. 1, pp. 55–64, 1993.