

Sensor fusion and State estimation of the robot

Francesco Nori & Silvio Traversaro

iCub Facility

Istituto Italiano di Tecnologia, Italy

Contact email: `name.surname@iit.it`

Maurice Fallon

School of Informatics

University of Edinburgh, UK

Contact email: `name.surname@ed.ac.uk`

August 10, 2016

Abstract

In this chapter we review previous estimation and sensor fusion approaches in the field of humanoid robotics. Focus is primarily but not exclusively on state estimation. We model humanoids as free-floating mechanical systems subject to external forces and constrained by whole-body distributed rigid contacts. Besides state estimation we consider a number of related problems including the estimation of joint torques, contact forces, contact positions and the center of mass. Previous approaches are discussed with reference to a number of features relevant in humanoids applications: computational complexity, observability analysis, modeling accuracy and floating base parameterization. The chapter concludes with a section on future directions and open problems, revisited under the light of previously discussed approaches.

1 Introduction

Within a control system, the roll of the state estimator is to combine, sometimes disparate, sensor measurements of the input and output of the system to create an overall estimate of the state (often including a measure of uncertainty) which is often then used as input to the controller. In humanoid robots the system state estimation is a challenging problem for the reasons hereafter discussed.

Differently from traditional industrial manipulators, humanoids are *free-floating systems* — as opposed to being rigidly fixed to the ground. The global pose of a humanoid therefore contains the overall position and orientation of the system, which are not directly sensed. Two complications therefore arise. Firstly, the parameterization of orientations cannot be simultaneously minimal and non-singular. To address this concern, previous approaches either considered a minimal representation (e.g. Euler angles) or a singularity-free parameterization (e.g. unit quaternions). Secondly, no direct sensor can estimate the free-floating component of the system pose (external motion capture systems aside). State estimators therefore need to rely on indirect measurements such as the kinematic constraints arising from footsteps making ground contact.

Humanoid motion controllers are often applied in *real-time* and it is often desirable to have computationally efficient implementations of their computations. For example, efficient solutions exist for solving the forward dynamics [13], i.e. for estimating the joint accelerations given the system state and the applied joint torques. Similarly, computationally efficient solutions exist for solving the inverse dynamics, i.e. for computing the joint torques given the system state and joint accelerations. Whole-body state estimation instead is a problem for which different dynamic models with different level of complexity have been considered in previous literature, trading-off between accuracy and computational complexity. Planar models (e.g. linear inverted pendulum) have been progressively replaced by articulated and three whole-body dimensional models. Dedicated observability analyses have been conducted to cope with the complexity and nonlinearity of these models. Even though joint positions are directly measured, velocities and accelerations require filtering which causes lags and therefore instability in the controller. Gyroscope and accelerometers can mitigate these drawbacks.

When in *contact with the environment*, humanoids are often represented as free-floating systems subject to physical constraints that represent the environmental contacts. Their state therefore is constrained to and evolves in a manifold whose topology is non-trivial, especially when the system is constrained at multiple contact locations. The state estimation problem should take these constraints into account to exclude those hypotheses which are not compatible

with the physical bounds of the considered system, including joint limits and collisions. Contacts and constraints generate forces normal to the constraint surface [23]. In this sense state estimation is crucial but a full characterization of the humanoid whole-body dynamics requires also the estimation of internal and external forces. Within this context, internal forces lead to joint torques and motor friction identification; external force allow estimation of exchanged contact forces with humans (for safety) and with the environment (for postural control).

Finally, estimation often relies on a several parameters which influence the estimation accuracy itself. Examples of these parameters, often referred to *hyper-parameters*, are sensor bias and sensor positioning. Tuning of these parameters can be either handled on-line with some adaptive strategies or off-line with suitably defined batch procedures. Sensor bias is often assumed to be a slowly drifting corrupting signal and it has been estimated on-line by assuming a random walk dynamics model. Other calibrations, such as the sensor positions, are typically estimated off-line with a batch procedure.

2 Structure

The present chapter surveys previous approaches to several estimation problems related to humanoid robotics. The chapter is divided into sections. Each section focuses on a specific estimation problem as described in Section 4 and previous approaches are categorized according to the estimated variables and estimation features. Table 1 summarizes the proposed classification. A first group of columns clusters previous approaches according to the type of estimated variables. State estimation is subdivided into joint related (i.e. joint positions, velocities and accelerations) and floating-base related quantities (i.e. floating-base position, orientation and accelerations). Force estimation is subdivided into internal (i.e. joint torques) and external (i.e. contact) force, the latter being subdivided into magnitude, direction and position estimation. Estimated variables also include the center of mass position and sensor biases, which are often included among the state variables. A second group of columns clusters previous approaches according to the features of the estimation algorithm. As discussed in the introduction, several features are considered. The column labeled “real-time performances” refers to the computational cost of the estimation procedure. Two columns concern hyper-parameter calibration and in particular sensor geometric calibration and joint friction identification. The column “observability analysis” identifies the approaches which propose a mathematical analysis to guarantee that estimated variables can be inferred from available sensors. Other important features are related to the trade-off between accuracy and complexity of the underlying dynamical model that represent the humanoid body structure. Within this context, we distinguish between three-dimensional versus planar and articulated versus rigid body models. The column “global parameterization” identifies those approaches which adopt a non-singular versus minimal parameterization for the free-floating state orientation. Finally, a dedicated column identifies previous works explicitly modeling physical constraints (e.g. joint limits and positivity constraints) in the estimation.

The chapter is organized as hereafter described. Section 3 introduces the notation used in the chapter. Section 4 describes the estimation problem and Section 5 discuss the current state of the art. Section 6 speculates on future directions and defines open problems. Finally Section 7 draw the conclusions.

3 Notation

3.1 Points and coordinate frames

A *frame* is defined as the combination of a point (called *origin*) and an *orientation frame* in 3D space. We typically employ a capital letter to indicate a frame. Given a frame A , we will indicate with o_A its origin and with $[A]$ its orientation frame. Formally, we write this as $A = (o_A, [A])$.

Frames can be time varying. They can be used, for example, to describe the position and orientation in space of a rigid body as time evolves. They are also used to expressed a coordinate system for a wrench exchanged by two bodies or used to define a coordinate system so as to describe a robot task (like a frame attached to the center of mass and oriented as the inertial frame).

Newtonian mechanics requires the definition of an *inertial* frame. In this document, we usually indicate this frame simply with A (the Absolute frame). As is common practice, for robots operating near the Earth surface, we will assume the frame A to be fixed to the world’s surface and disregard non-inertial effects due to the Earth’s motion.

Section	Estimated variables									Estimation features									
	Joint-positions	Joint-velocities	Joint-accelerations	Floating-base position	Floating-base velocity	Floating-base accelerations	Joint-torques	External forces magn. and dir.	External forces position	Center of mass position	Sensor bias	Real-time performance	Joint friction modeling	Sensor geometric calibration	Observability analysis	Three-dimensional model	Articulated-body model		Global parametrization
5.1	✓	✓	✓								✓	✓		✓					Vihonen et al. [41] Xinjilefu et al. [47] Rotella et al. [32]
5.2							✓					✓	✓	✓					De Luca et al. [7] Tien and Albu-Schäffer [39] Zhang et al. [49] Park and Chung [29] Sotoudehnejad et al. [35] Fumagalli et al. [14]
5.3.2							✓	✓	✓			✓							Park and Khatib [28] Del Prete et al. [9]
5.3.1							✓	✓	✓	✓				✓					Del Prete et al. [8] Ciliberto et al. [6] Magrini et al. [20] Manuelli and Tedrake [21] Petrovskaya et al. [30]
5.4	✓	✓		✓	✓				✓	✓	✓	✓			✓	×	×		Stephens [37] Xinjilefu and Atkeson [43] Aoustin et al. [2] Benallegue and Lamiroux [3] Pongsak et al. [31] Rotella et al. [33] Xinjilefu et al. [44] Xinjilefu et al. [45]
5.5																			Thompson and Kagami [38] Chitta et al. [5] Alcantarilla et al. [1] Yoon et al. [48] Gonzalez-Aguirre et al. [15] Fallon et al. [11] Oriolo et al. [27]
5.6	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓		✓		✓	✓	×	Nori et al. [25] Lowrey et al. [19] Wu et al. [42]

Table 1: The table presents a classification of previous works on state estimation for humanoid robots. The adopted classification criteria are subdivided into estimated variables and estimation features. Estimated variables are separated into joint and floating base variables. Additional relevant variables have been included in the classification: joint torques, external forces and center of mass position.

3.2 Coordinate vector of a point

Given a point p , its coordinates with respect to a frame $A = (o_A, [A])$ are collected in the *coordinate vector* ${}^A p$. The coordinate vector ${}^A p$ represents the coordinates of the 3D geometric vector $\vec{r}_{o_A, p}$ connecting the origin of frame A with the point p , pointing towards p , expressed in the orientation frame $[A]$, that is

$${}^A p := \begin{bmatrix} \vec{r}_{o_A, p} \cdot \vec{x}_A \\ \vec{r}_{o_A, p} \cdot \vec{y}_A \\ \vec{r}_{o_A, p} \cdot \vec{z}_A \end{bmatrix} \in \mathbb{R}^3, \quad (1)$$

with $\cdot$ denoting the scalar product between two vectors and $\vec{x}_A, \vec{y}_A, \vec{z}_A$, the unit vectors defining the orientation frame $[A]$.

3.3 Change of orientation frame

Given two frames A and B , we will employ the notation:

$${}^A R_B \in \text{SO}(3) \quad (2)$$

to denote the coordinate transformation from frame B to frame A . The coordinate transformation ${}^A R_B$ only depends on the relative orientation between the orientation frames $[A]$ and $[B]$, irrespectively of the position of the origins o_A and o_B .

3.4 Homogeneous transformation

To describe the position and orientation of a frame B with respect to another frame A , we employ the 4×4 homogeneous matrix

$${}^A H_B := \begin{bmatrix} {}^A R_B & {}^A o_B \\ 0_{1 \times 3} & 1 \end{bmatrix}. \quad (3)$$

Given a point p , the homogeneous transformation matrix ${}^A H_B$ can be also used to map the coordinate vector ${}^A p$ to ${}^B p$ as follows. Let ${}^A \bar{p}$ and ${}^B \bar{p}$ denote the *homogenous representation* of ${}^A p$ and ${}^B p$, respectively. That is, let ${}^A \bar{p} := ({}^A p; 1) \in \mathbb{R}^4$ and likewise for ${}^B \bar{p}$ (note that $;$ indicates row concatenation). Then

$${}^A \bar{p} = {}^A H_B {}^B \bar{p}, \quad (4)$$

which is the matrix form of ${}^A p = {}^A R_B {}^B p + {}^A o_B$. We refer to [23, Chapter 2] for further details on homogeneous representation of rigid transformations.

3.5 Velocity vectors (twists)

In the following, given a point p and a frame A , we define

$${}^A \dot{p} := \frac{d}{dt} ({}^A p). \quad (5)$$

In particular, when p is the origin of a frame, e.g., $p = o_B$, we have

$${}^A \dot{o}_B = \frac{d}{dt} ({}^A o_B).$$

It is important to note that, by itself, expressions like $\dot{o}_B$ or $\dot{p}$ have *no* meaning. Similarly to (5), we also define

$${}^A \dot{R}_B := \frac{d}{dt} ({}^A R_B) \quad (6)$$

and

$${}^A \dot{H}_B := \frac{d}{dt} ({}^A H_B) = \begin{bmatrix} {}^A \dot{R}_B & {}^A \dot{o}_B \\ 0_{1 \times 3} & 0 \end{bmatrix}. \quad (7)$$

The relative velocity between a frame B with respect to a frame A can be represented by the time derivative of the homogenous matrix ${}^A H_B \in \text{SE}(3)$. A more compact representation of ${}^A \dot{H}_B$ can be obtained multiplying it by the inverse of ${}^A H_B$ on the left or on the right. In both cases, the result is an element of $\mathfrak{se}(3)$ that will be called a *twist*. Premultiplying on the left, one obtains

$$\begin{aligned} {}^A H_B^{-1} {}^A \dot{H}_B &= \begin{bmatrix} {}^A R_B^T & -{}^A R_B^T {}^A o_B \\ 0_{1 \times 3} & 1 \end{bmatrix} \begin{bmatrix} {}^A \dot{R}_B & {}^A \dot{o}_B \\ 0_{1 \times 3} & 0 \end{bmatrix} \\ &= \begin{bmatrix} {}^A R_B^T {}^A \dot{R}_B & {}^A R_B^T {}^A \dot{o}_B \\ 0_{1 \times 3} & 0 \end{bmatrix}. \end{aligned} \quad (8)$$

Note that ${}^A R_B^T {}^A \dot{R}_B$ appearing on the right hand side of (8) is skew symmetric. Define ${}^B v_{A,B}$ and ${}^B \omega_{A,B} \in \mathbb{R}^3$ so that

$${}^B v_{A,B} := {}^A R_B^T {}^A \dot{o}_B, \quad (9)$$

$${}^B \omega_{A,B}^\wedge := {}^A R_B^T {}^A \dot{R}_B. \quad (10)$$

The *left trivialized* velocity of frame B with respect to frame A is

$${}^B \mathbf{v}_{A,B} := \begin{bmatrix} {}^B v_{A,B} \\ {}^B \omega_{A,B}^\wedge \end{bmatrix} \in \mathbb{R}^6. \quad (11)$$

By construction,

$${}^B \mathbf{v}_{A,B}^\wedge = {}^A H_B^{-1} {}^A \dot{H}_B. \quad (12)$$

Note the slight abuse of notation in using the hat operator $\wedge$ in (10) and (12) that maps a vector into its corresponding matrix representation (respectively, from $\mathbb{R}^3$ to $\mathbb{R}^{3 \times 3}$ and from $\mathbb{R}^6$ to $\mathbb{R}^{4 \times 4}$).

Specularly to what done in (8), right multiplying ${}^A \dot{H}_B$ by the inverse of ${}^A H_B$ leads to

$$\begin{aligned} {}^A \dot{H}_B {}^A H_B^{-1} &= \begin{bmatrix} {}^A \dot{R}_B & {}^A \dot{o}_B \\ 0_{1 \times 3} & 0 \end{bmatrix} \begin{bmatrix} {}^A R_B^T & -{}^A R_B^T {}^A o_B \\ 0_{1 \times 3} & 1 \end{bmatrix} \\ &= \begin{bmatrix} {}^A \dot{R}_B {}^A R_B^T & {}^A \dot{o}_B - {}^A \dot{R}_B {}^A R_B^T {}^A o_B \\ 0_{1 \times 3} & 0 \end{bmatrix}. \end{aligned} \quad (13)$$

Define ${}^A v_{A,B}$ and ${}^A \omega_{A,B} \in \mathbb{R}^3$ as

$${}^A v_{A,B} := {}^A \dot{o}_B - {}^A \dot{R}_B {}^A R_B^T {}^A o_B \quad (14)$$

$${}^A \omega_{A,B}^\wedge := {}^A \dot{R}_B {}^A R_B^T. \quad (15)$$

The *right trivialized* velocity of B with respect to A is then defined as

$${}^A \mathbf{v}_{A,B} := \begin{bmatrix} {}^A v_{A,B} \\ {}^A \omega_{A,B}^\wedge \end{bmatrix} \in \mathbb{R}^6. \quad (16)$$

By construction,

$${}^A \mathbf{v}_{A,B}^\wedge = {}^A \dot{H}_B {}^A H_B^{-1}. \quad (17)$$

3.5.1 Expressing a twist with respect to an arbitrary frame

Straightforward algebraic calculations allow to show that the right and left trivialized velocities ${}^A \mathbf{v}_{A,B}$ and ${}^B \mathbf{v}_{A,B}$ are related via a linear transformation. Following the notation introduced in Featherstone [13], we denote this transformation ${}^A X_B$ and define it as

$${}^A X_B := \begin{bmatrix} {}^A R_B & {}^A o_B^\wedge {}^A R_B \\ 0_{3 \times 3} & {}^A R_B \end{bmatrix} \in \mathbb{R}^{6 \times 6}. \quad (18)$$

We have therefore

$${}^A \mathbf{v}_{A,B} = {}^A X_B {}^B \mathbf{v}_{A,B}. \quad (19)$$

The inverse transformation is simply given by ${}^B X_A = {}^A X_B^{-1}$, an exercise we leave to the reader (recall that ${}^A o_B = -{}^A R_B {}^B o_A$).

3.6 Jacobians

The goal of this section is to arrive at a precise, unambiguous notation to denote the ‘geometric Jacobians’ for fixed-based and, in particular, free-floating multibody systems.

In this section, A will denote the (absolute) *inertial frame* and B a frame rigidly attached to one of the bodies composing the multibody system, selected to be the *floating base*. The configuration of a free-floating multibody system will be parametrized using $q = (H, q_J) \in \text{SE}(3) \times \mathbb{R}^{n_J}$, with $H = {}^A H_B \in \text{SE}(3)$ representing the pose (position and orientation) of the floating base frame B and $q_J \in \mathbb{R}^{n_J}$ the internal joint displacements. The configuration space (more correctly, the configuration manifold) has correspondingly dimension $n = 6 + n_J$.

Let E be a frame (rigidly) attached to an arbitrary body to be used, e.g., for the specification of a task to be executed by the robot. The frame E could represent, e.g., the pose of a specific frame rigidly attached to a *end effector* of a robot manipulator, like a hand or foot of a humanoid robot. Let

$${}^A H_E = {}^A H_E(q) = {}^A H_E(H, q_J) \quad (20)$$

denote the homogeneous transformation expressing E with respect to A as a function of the configuration $q = (H, q_J)$.

Let δH denote an infinitesimal perturbation of the pose of the floating base –in the language of differential geometry, $\delta H \in T_H \text{SE}(3)$ – and δq_J an infinitesimal perturbation of the joint displacements. Then, the corresponding infinitesimal perturbation of frame E can be computed as

$${}^A \delta H_E = {}^A D_1 H_E(H, q_J) \cdot \delta H + {}^A D_2 H_E(H, q_J) \cdot \delta q_J, \quad (21)$$

where we recall that H is a short form for ${}^A H_B$ and δH a short form for ${}^A \delta H_B$. Define the trivialized infinitesimal perturbations ${}^E \Delta_{A,E}$ and ${}^B \Delta_{A,B} \in \mathbb{R}^6$ such that

$${}^E \Delta_{A,E}^\wedge = {}^A H_E^{-1} {}^A \delta H_E \quad (22)$$

and

$${}^B \Delta_{A,B}^\wedge = {}^A H_B^{-1} {}^A \delta H_B. \quad (23)$$

Combining (22) and (23) together with (21), it is straightforward to show that ${}^E \Delta_{A,E}$ depends linearly on ${}^B \Delta_{A,B}$ and δq_J . Such a linear map defines the “geometric Jacobian” for a floating base system. It will be indicated with the symbol ${}^E J_{A,E/B}$ to indicate that the Jacobian allows to compute the infinitesimal perturbation of frame E relative to frame A , expressed with respect to frame E (a first left-trivialization), based on the infinitesimal perturbation of the internal joint configuration and that of the floating base B , the latest expressed with respect B (a second left-trivialization). For this reason, we will refer to ${}^E J_{A,E/B}$ as a *double left-trivialized* geometric Jacobian. In formulas, we get

$${}^E \Delta_{A,E} = {}^E J_{A,E/B}({}^A H_B, q_J) \begin{bmatrix} {}^B \Delta_{A,B} \\ \delta q_J \end{bmatrix}. \quad (24)$$

The infinitesimal perturbation of frames E and B can be clearly be expressed with respect to arbitrary frames, let us say, C and D . The geometric Jacobian ${}^D J_{A,E/C}$ is related to ${}^E J_{A,E/B}$ by the transformation

$${}^D J_{A,E/C} = {}^D X_E {}^E J_{A,E/B} {}^B Y_C, \quad (25)$$

with

$${}^B Y_C := \begin{bmatrix} {}^B X_C & 0_{6 \times n_J} \\ 0_{n_J \times 6} & I_{n_J \times n_J} \end{bmatrix}. \quad (26)$$

The special case ${}^{E[A]} J_{A,E/B[A]}$, obtained from the above formula with $D = E[A]$ and $C = B[A]$, will be denoted the *mixed-mixed* geometric Jacobian.

3.7 Free-floating rigid body dynamics

Here we introduce the notation that we use to write, in a compact form, the equations of motion of a free-floating rigid body system. As mentioned in Section 3.6, the configuration of a free-floating rigid body system will be parametrized as $q = (H, q_J) := ({}^A H_B, q_J)$, where A is the inertial frame, B the (selected) floating base frame and $q_J \in \mathbb{R}^{n_J}$ the joint displacements.

The kinematics of the (floating base) rigid body system are written treating the configuration manifold $Q = \text{SE}(3) \times \mathbb{R}^{n_J}$ as the Lie group defined by the group direct product $\text{SE}(3) \times \mathbb{R}^{n_J}$. Seeing Q as a Lie group, employing the mixed velocity of the floating based B , allows one to write the derivative of the configuration $\dot{q} = (\dot{H}, \dot{q}_J)$ as

$$\dot{q} = q Y \nu, \quad (27)$$

where $\nu = (v, \dot{q}_J)$ denotes the *mixed generalized velocity* of the floating-base rigid-body system, i.e., $\nu := {}^{B[A]} \nu = ({}^{B[A]} v_{A,B}, \dot{q}_J)$, and $Y := {}^B Y_{B[A]}$ is the natural extension of the velocity transformation X , introduced in Section 3.5.1, to the Lie algebra of Q . The structure of the generalized velocity transformation Y has been introduced previously in (26) and it is formally defined in such a way that

$$\begin{aligned} Y \nu &= {}^B Y_{B[A]} {}^{B[A]} \nu \\ &= {}^B Y_{B[A]} ({}^{B[A]} v_{A,B}, \dot{q}_J) \\ &= ({}^B X_{B[A]} {}^{B[A]} v_{A,B}, \dot{q}_J) \\ &= ({}^B v_{A,B}, \dot{q}_J) =: {}^B \nu. \end{aligned} \quad (28)$$

Note how Y acts on the floating base velocity v , leaving the joint velocities $\dot{q}_J$ unaltered. The configuration q and the mixed generalized velocity ν can then be used to write the dynamics of the floating-based rigid-body system as

$$M(q) \dot{\nu} + C(q, \nu) \nu + G(q) = S^\top \tau + \sum_{i \in \mathcal{I}_N} J_i^T(q) f_i \quad (29)$$

with M the (mixed) mass matrix, C the (mixed) Coriolis matrix, G the (mixed) potential force vector, $S := [0_{6 \times n_J}; I_{n_J \times n_J}]$ the *joint selection matrix*, τ the joint torques, $\mathcal{I}_N$ the set of closed contacts, $f_i := {}^{C_i[A]} f_i$ the i -th contact wrench, and

$$J_i(q) := {}^{C_i[A]} J_{A, L_i / B[A]}(q) \quad (30)$$

the mixed-mixed geometric Jacobian associated to the contact frame C_i , where we used the notation introduced in Section 3.6, where L_i is the frame *rigidly attached* to the link that is experiencing the i -th contact.

3.8 Additional definitions

Kinematic trees ordering. In the following sections, we restrict the analysis to robots described by kinematic trees (i.e. robots with no kinematic loops) composed of N_B rigid links numbered from 0 to N_B , 0 being the selected fixed base (world reference frame) and 1 being the selected floating base (reference rigid body in the articulated chain). Link numbering is obtained by defining a suitable spanning tree where each rigid link is associated to a unique node in the tree. Numbers can be always selected in a topological order so that each node i has a higher number than its unique parent $\lambda(i)$ and a smaller number than all the nodes in the set of its children $\mu(i)$. Links i and $\lambda(i)$ are coupled with the joint i whose joint motion constraints are modelled with $S_i \in \mathbb{R}^{6 \times n_i}$, n_i being the number of degrees of freedom of joint i and $n = n_1 + \dots + n_{N_B}$ the total number of degrees of freedom excluding the floating base.

Proper acceleration. Adopting a naming convection common in relative relativity theory but without entering the details of its definition, we will use the term proper acceleration. In the scope of this chapter, it is sufficient to identify the proper acceleration with the quantity measured by an accelerometer. Let o_S be the origin of the accelerometer frame S and let A be an inertial frame. In a gravity field g , the proper acceleration is ${}^S R_A ({}^A \ddot{o}_S - {}^A g)$. The proper acceleration of an object which is not accelerating in an inertial frame is $-g$; this is the case of a stationary accelerometer placed on a table. Moreover, the proper acceleration of a free-falling object is zero.

Wrench. We will often refer to the term wrench which corresponds to the six dimensional vector obtained by the vertical concatenation of a three dimensional force and a three dimensional torque.

4 Problem formulation

In this chapter we consider the estimation problems arising in the field of humanoid robotics. In controlling a system it is often assumed that the system pose is known. In the context of control theory, this requirement can be formalized by saying that a controller often requires an estimation of the system state. In this chapter we consider the state estimation problem for a particular class of systems, nominally free-floating articulated rigid-bodies. These systems belong to the broader class of lumped parameter mechanisms, for which the state has two components: the configuration and the associated velocity. We specifically focus on the problem arising from the fact that humanoids are free-floating systems. The robot pose is parametrized by a global set of coordinates. Since the pose includes the orientation of the base frame choosing a global parametrization implies that a non-minimal set of coordinates (i.e. homogenous transformations) has been preferred to minimal ones (e.g. Euler angles). The robot velocity ν instead includes the twist of the base from motion and therefore the robot velocity ν does not coincide with the time derivative of the robot pose q . The algebraic constraints acting on the system which can be used to represent whole-body distributed rigid contacts. The system dynamics are therefore described by the following differential algebraic equation:

$$q = (H, q_J), \quad \nu = (\dot{v}, \dot{q}_J), \quad (31a)$$

$$M(q)\dot{\nu} + C(q, \nu)\nu + G(q) - \sum_{i \in \mathcal{I}_N} J_i^T(q) f_i - \sum_{i \in \mathcal{I}_N^x} J_{i,x}^T(q) f_i^x = S^\top \tau, \quad (31b)$$

$$\forall i \in \mathcal{I}_N \quad J_i(q)\dot{\nu} + \dot{J}_i(q, \nu)\nu = 0. \quad (31c)$$

The algebraic constraints (31c) represent the rigid contacts acting on the free-floating system (e.g. foot contacts). The differential equation (31b) corresponds to the free-floating system (29) with additional contact forces f_i ($i \in \mathcal{I}_N$) and external forces f_i^x ($i \in \mathcal{I}_N^x$). The location of the latter is identified by the associated Jacobian $J_{i,x}(q)$. A peculiarity of free-floating systems is that the system configuration (q) has two components: the joint positions ($q_J \in \mathbb{R}^{n_J}$ in Section 3.7) and the system rotation-translation ($H \in SE(3)$ in Section 3.7). In this chapter we consider different estimation problems associated with humanoid robots described as in (31).

Problem 1 (Joint velocity and acceleration estimation). *Consider the differential algebraic equation (31). Given a set of sensors estimate the joint velocities $\dot{q}_J \in \mathbb{R}^{n_J}$ and acceleration $\ddot{q}_J \in \mathbb{R}^{n_J}$.*

Problem 2 (Joint torques estimation). *Consider the differential algebraic equation (31). Given a set of sensors estimate the joint torques $\tau \in \mathbb{R}^{n_J}$.*

Problem 3 (External force estimation). *Consider the differential algebraic equation (31). Given a set of sensors estimate the external forces f_i ($i \in \mathcal{I}_N$), f_i^x ($i \in \mathcal{I}_N^x$) and their locations $J_i(q)$, $J_{i,x}(q)$.*

Problem 4 (Floating-base estimation). *Consider the differential algebraic equation (31). Given a set of sensors estimate the floating-base component $H \in SE(3)$ of the state vector q .*

5 Problem solution: state of the art

In this section we propose a number of estimation problems which have been addressed in the field of humanoid robotics. Section 5.1 presents previous literature on the problem of joint velocity and acceleration estimation (Problem 1). Section 5.2 presents previous literature on estimating joint velocities and accelerations (Problem 2). Section 5.2 focuses instead on the problem of estimating joint torques (Problem 2). The problem of estimating external and contact forces (Problem 3) is considered in Section 5.3, which considers separately the problems of force localization (Section 5.3.2) and magnitude, direction estimation (Section 5.3.1). Two dedicated sections present solutions of the floating-base estimation problem (Problem 4) subdivided into two subcategories. In particular, Section 5.4 and Section 5.5 discuss estimations which rely or not on range and vision sensors. Finally, Section 5.6 considers combined approaches which handle simultaneously different estimation problems (e.g. calibration, physical consistency, torque and floating base estimations).

5.1 Joint velocity and acceleration estimation

Classical control algorithms for fixed-base robots [34] and their extensions to free-floating robots [26] require access to joint velocity measurements. Traditionally, joint positions q_J can be directly measured with joint encoders from which joint velocities $\dot{q}_J$ and accelerations $\ddot{q}_J$ are computed via numerical differentiation. These estimations are often too noisy to be used in a control algorithm and filtering them introduces a potentially-destabilizing time delay, preventing the use of feedback gains high enough to achieve satisfactory stiffness and damping. Recently the use of miniaturized gyroscopes and accelerometers have been proposed as a possible source of information to achieve reliable velocity and acceleration estimation. In the following we discuss three approaches recently proposed by Vihonen et al. [41], Xinjilefu et al. [47] and Rotella et al. [32] to exploit whole-body distributed accelerometers and gyroscopes.

Vihonen et al. [41] addressed the problem of joint velocity and acceleration estimation to circumvent issues related to limited joint-encoder bandwidth and lag induced by classical filtering. The authors propose an estimation strategy based on rate gyros and accelerometers only, pointing out that magnetometers are unusable in many practical situations. Estimated quantities include joint velocities and joint accelerations. Joint velocities are recursively estimated from the chain root to the leaves (see the link ordering defined in Section 3.8) as the difference between the link angular velocity (as measured by the attached gyroscope) and the sum of previous joint velocities. Similarly, joint accelerations are estimated as the difference between the link angular acceleration and the sum of previous joint accelerations. Since no direct link acceleration measurement is available, the authors propose two estimation strategies: numerical differentiation on one hand, direct estimation on the other. The latter is based on six linear accelerometers rigidly attached at different locations of a rigid link and authors show that the link angular acceleration can be estimated by combining the measurements from accelerometers and gyroscopes. The authors present some accuracy results: angular error $0.7[deg]$ peak absolute error (PAE) and $0.16[deg]$ root mean square error (RMSE), joint velocity error $0.07[deg/s]$ RMSE, joint acceleration error $8.3[deg/s^2]$ RMSE.

Rotella et al. [32] proposed a very similar approach with applications to humanoid robotics. Joint accelerations are estimated with a procedure similar to the one presented in [41] which exploits a second inertial measurement unit attached to each link. Authors propose a calibration procedure which allows estimation of the sensors' positions and the orientations. The procedure relies on the gyroscopes measurements whose time derivatives are numerically computed with zero-delay filters, exploiting the fact that off-line calibration procedures can rely on acausal estimations strategies. Accuracy of the estimation is improved with a Kalman filter to fuse joint velocities and accelerations computed from inertial sensors with joint position sensor measurements. The joint velocity filter adopts two different definitions of its state dynamics. First, authors propose to use measured joint torques and end-effector wrenches with the robot's whole-body dynamic model. Second, authors proposed a simple first order approximation of joint velocities by assuming that joint accelerations are either estimated or replaced by the desired or predicted joint accelerations from the applied controller. Experimental results have been proposed to compare link 3D velocities, with joint velocities and calibrated joint velocities; the validation is conducted by proving that estimated velocities allow to increase velocity gains with improved tracking performance in performing a sinusoid tracking experiment.

Xinjilefu et al. [47] proposed a similar approach based on a Kalman filter with a state containing the vector of joint velocities. The measurement set included the derivative of the sensed joint positions and the available gyroscopes. As in [32], the state dynamics include desired acceleration computed by the applied controller (in this case a QP controller). A sequential Kalman filter is adopted to reduce computational complexity with a block diagonal noise covariance matrix. Gyroscopes orientation are calibrated in $SO(3)$ with an analytic solution of the associated least squares problem identical to the one proposed in [32]. Gyroscope biases are calibrated at the beginning of the experiment and no results of on-line estimation were provided. The comparison of the proposed Kalman filter with classical (causal and acausal) filtering strategies shows that the Kalman filter traces are about $15\text{-}20[ms]$ faster than the low pass filtered ones.

5.2 Joint torque estimation

Joint torques are yet another important component of the differential equation (31) governing a humanoid dynamic behavior. Even though joint torques are not part of the system state, their estimation is crucial to improve the system state estimation accuracy. In this section we discuss the problem of joint torque estimation which often boils down into the problem of distinguishing the two components which contribute to the total joint torque: the actuating motor torques, τ , and the external joint torques due to the i -th contact, $J_i^T(q)f_i$ in (31b). In the following we discuss previous approaches to the problem of joint torque estimation. De Luca et al. [7], Tien and Albu-Schäffer [39], Park and

Chung [29] and Sotoudehnejad et al. [35] have proposed approaches which estimate external joint torques given a measurements of the joint motor torques. Zhang et al. [49] and Fumagalli et al. [14] have focused on the problem of estimating motor torques given total joint torques.

De Luca et al. [7] proposed an approach to estimate the effect of external forces on joint torques from motor torque measurements. Their approach is based on the definition of a dynamic quantity, the residual, which acts as a collision identification signal. Its has two main features: first, it can be computed from the motor torques and the joint positions and velocities; second, its dynamics are governed by a linear differential equation and therefore the residual is zero if no external contact force is applied to the manipulator. The definition of the residual for both rigid and elastic joint manipulators is proposed by maintaining the two aforementioned features. Tien and Albu-Schäffer [39] adopted the approach in [7] at the actuator level and exploited the residual computations to estimate the joint friction, modeled as the sum of two linear terms: Coulomb and viscous friction.

Park and Chung [29] extended the residual approach proposed by De Luca et al. [7] to estimate both actuating motor torque and external link torque. Authors, point out that the residual definition in [7] can be framed within a general theoretical framework, named linear disturbance observer (DOB). The proposed approach is tested on an elastic joint prototype with custom torque sensors.

Sotoudehnejad et al. [35] proposed yet another extension of the residuals concept, focusing in particular on the problem of using time varying thresholds on the residuals to detect external contacts. Authors proposed an accurate model of different modeling errors to define tight thresholds. Considered errors include friction models and uncertainties on the manipulator dynamic parameters (mass, inertias, centers of mass).

Zhang et al. [49] proposed a technique to estimate joint torques by accurate measurements of the harmonic drive flex spline deflections. The proposed experiments rely on two encoders whose difference is proportional to the spline deflection. An high resolution encoder is placed on the joint and a standard optical incremental encoder measures the motor position. The obtained estimation accuracy is satisfactory, considering that the root mean square error averages $0.25[Nm]$ and never exceeds $0.5[Nm]$. This level of accuracy requires extensive modeling of phenomena which are often neglected. First authors characterized the harmonic drive kinematic errors (the deviation from the ideal gearbox behavior) with periodic functions. Second, authors modeled the flex-spline stiffness with a quadratic curve subject to an hysteresis effect.

Fumagalli et al. [14] considered the problem of joint torques estimation with whole-body distributed force/torque and tactile sensors. The proposed estimation strategy consists in a modified recursive Newton-Euler algorithm (RNEA) for solving the associated inverse dynamics problem. The classical RNEA boundary conditions (end-effector forces and base accelerations) are replaced with available measurements (embedded force/torque sensors and inertial measurement units). The RNEA propagation ordering is redefined according to the position of external contacts, sensed at real-time by the tactile sensors. Successive publications from the authors have shown that the proposed estimation strategy can be used to estimate the joint viscous and Coulomb frictions and to implement inverse dynamic control strategy with torque control [26]. Moreover similar estimation strategies have been proven to be applicable on the HRP-2 humanoid, which is not equipped with tactile sensors and is not equipped with joint torque sensing technologies [10].

5.3 Force estimation

External f_i^x and contact f_i forces act as external inputs on humanoids. Their estimation is therefore fundamental for a complete characterization of the robot dynamic behavior. In the absence of reliable tactile technologies, previous approaches mainly focused on the problem of estimating the position of externally applied forces (Section 5.3.1). Within this context, authors proposed some strategies to estimate the applied force magnitude but all proposed solutions suffer of identifiability issues [20, 21]. Section 5.3.2 discusses exact solutions to the problem of estimating contact wrench magnitude and direction, relying on the assumption that contact locations are sensed with a tactile sensor.

5.3.1 Force position estimation

In this section we discuss previous approaches on the estimation of contact location which reduces to the problem of retrieving the Jacobian of the associated contact point.

Del Prete et al. [8] considered the problem of estimating the position of an external contact assuming that the contact generates a pure force (i.e. null torque) and its effects are sensed with a force/torque sensor embedded in the robot structure (e.g. at the robot base). The estimation relies on the classical force transformation principle [34,

Section 3.8.3] according to which the spatial transformation of a pure force generates a torque equal to the cross product of the application vector and the applied force. Estimating the application vector (i.e. the contact location) boils down to a linear least squares which identifies a unique solution if at least two non collinear forces. The proposed procedure allows the calibration of a tactile array. The achieved spatial resolution is $7mm$ but can be improved to $6mm$ projecting the estimated locations on the mesh representing the robot geometry.

Magrini et al. [20] proposed a contact estimation strategy which is based on the residual concept proposed by De Luca et al. [7]. The estimation of the contact location relies on a Kinect to detect the human posture and on a robot three dimensional rendering to predict possible contact location on the robot surfaces. Given the contact location, a least-square estimation is proposed to obtain the contact force magnitude and direction via a Moore-Penrose pseudo-inverse computation. Only the force components which do not lie on null space of the contact Jacobian can be estimated and this is a limitation of the proposed approach.

Manuelli and Tedrake [21] formulated another contact estimation strategy based on [7] and extended it to the case of multiple contacts. Authors propose to solve two problems simultaneously: contact location on the one hand and external force estimation on the other. To reduce the associated computational complexity, the two problems are separated to exploit the fact that the force estimation problem is convex if contact locations are known. The proposed estimation problem is solved with a particle filter. Authors interestingly highlight several limitations of the proposed approach, focusing in particular on the effect of model uncertainties and on force magnitude identifiability issues.

5.3.2 Force magnitude and direction estimation

In this section we focus on the problem of contact wrench estimation. Previous approaches assume that contacts happen at known locations [28, 30] which might be sensed with a whole-body distributed tactile sensor [9, 6].

Park and Khatib [28] considered the problem of estimating contact forces assuming a spring contact model for the external forces. A classical task-space inverse dynamics controllers is adopted to control contact forces. The proposed estimation technique, named Active Observer (AOB), consists of a Kalman filter for estimating the difference between the input command force and measured output force.

Petrovskaya et al. [30] extended the AOB estimator in [28] to include the geometric model of the interaction. In particular, the proposed approach defines a probabilistic model of the robot and the environment. The estimation strategy aims at estimating the position of the contact point on the robot, while maintaining constant the model of the environment. Estimation is limited to the single contact case even if experiments include some multi-contact scenarios handled with compliant exploration strategy described by Park and Khatib [28].

Del Prete et al. [9] proposed a strategy to estimate contact wrenches given whole-body distributed force/torque and tactile sensors. The estimation strategy relies on the joint torque estimations described by Fumagalli et al. [14] and detailed in Section 5.2. As a by-product of the rearranged Newton-Euler recursion, authors present an algorithm to compute the total (external) wrench acting on a subpart. Subparts are defined by the kinematic subchains obtained by dividing the robot at the level of the available force/torque sensors. Exact estimation is possible only if there is exactly one contact wrench per subchain; otherwise an approximated solution based on linear least-squares is proposed.

Ciliberto et al. [6] extended previous approaches by Del Prete et al. [9, 8] which rely on distributed tactile sensors. At first, the approach in [8] is used to perform a spatial calibration of the tactile elements. Then, another calibration procedure is used to estimate the tactile sensor stiffness by correlating the force/torque measurements with the perceived tactile pressure. The resulting stiffness calibration allows to retrieve the normal component of all forces applied on the tactile sensors thus reducing the approximation errors described in [9].

5.4 Free-floating state estimation without range and vision sensors

In this section we consider the problem of estimating the state of a free-floating mechanical system. The section is divided into two categories depending on the parameterization adopted for the free-floating system. Section 5.4.1 presents approaches that consider the center of mass position as part of the system state. In particular Xinjilefu and Atkeson [43], Stephens [37] and Pongsak et al. [31] proposed approaches focusing on the estimation of the center of mass position and velocity. For sake of reducing computational complexity, these approaches have dealt either with a simplified model of the system (e.g. linear inverted pendulum model, LIPM) or with a reduced (if not minimal) set of coordinates (e.g. joint coordinates). Section 5.4.2 presents previous approaches which parametrized the free-floating system state with the floating base pose, represented by the quantities $H \in SE(3)$ and $v \in \mathbb{R}^3$ in (31). These approaches typically define a full-body articulated model. Benallegue and Lamiraux [3] focused on estimating

the humanoid pose assuming that uncertainty is concentrated at the ankle. Rotella et al. [33] included the humanoid odometry in the pose estimation, considering the kinematic constraints induced by footholds. Xinjilefu et al. [44, 45] proposed two approaches to reduce the computational complexity of whole-body estimation and to include linear inequalities to the estimation.

5.4.1 Center of mass estimation

Pongsak et al. [31] proposed one of the very first approaches to the problem of estimating the center of mass position and velocity of a humanoid robot. Considered measurements are provided by two linear accelerometers and a gyroscope. The estimation strategy considers the humanoid as a single rigid body subject to an external wrench, represented as a measured input in the rigid body dynamics. The resulting dynamical system is linearized about the horizontal and stable position. All measurements (accelerometers and gyroscope) are corrupted by noises. Modeling these noises with the information provided by the sensor manufacturers, authors propose an optimal H_2 -filter for the associated estimation problem.

Stephens [37] has proposed a linear inverted pendulum model (LIPM) to estimate some kinematic (centre of mass position and velocity) and some dynamic (centre of pressure and external forces) variables. The proposed estimation strategy deals with modeling errors such as measurement offsets and external disturbances. The author consider the associated state equations (e.g. center of mass offset, external force offset) and conduct an observability analysis on each of the proposed alternatives. The comparison of the different approaches is conducted with a validation experiment on the Sarcos humanoid. The drawn conclusion is that the presented models will not allow the estimation of both center of mass offset and external force disturbances and that there is a trade-off between precisely estimating the disturbance and the state.

Xinjilefu and Atkeson [43] proposed yet another approach to the problem of estimating the centre of mass position and velocity of a walking humanoid robot. Goal of the paper is to quantitatively compare the results obtained with the approach proposed in [37] with those obtained with an alternative estimation strategy based on a five-links humanoid planar model. In the linear inverted pendulum model, considered sensors are the feet position with respect to the center of mass and the distribution of vertical forces among feet; in the five-links model sensors are joint angles, joint torques, force/torque sensors at contacts, and an IMU mounted on the torso. The estimation is conducted over several footsteps and virtual measurements are included in the estimation to represent non-slipping feet contacts and the associated kinematic constraints. Experiments show that the Kalman filter based on the planar five-links model outperforms the LIPM filter, mainly due to the torso rotating around the hip joint during walking.

5.4.2 Floating-base state estimation

In the previous section we discussed approaches which have been proposed in literature to estimate position, velocity and acceleration of free-floating articulated rigid bodies. These approaches considered either simplified models or reduced (if not minimal) set of coordinates (e.g. joint coordinates) for sake of reducing computational complexity. In this section we discuss recent approaches have considered full-body models which guarantee the physical consistency of the resulting estimation.

Aoustin et al. [2] considered a very specific estimation problem: the orientation of the planar humanoid Rabbit. The particularity of the proposed approach is the conducted observability analysis which relies on the existence of a state-space transformation of the underlying nonlinear dynamic model. The transformation leads to an almost-linear representation (Brunovsky's form) which allows the application of different observers: a high gain observer, a step-by-step sliding mode observer, and an observer based on high order sliding mode differentiation. No significant differences were observed among the proposed estimators which lead to similar errors: less than $2[deg]$ error was observed on orientations and less than $0.65[rad/s]$ on angular velocities.

Benallegue and Lamiraux [3] considered a simplified floating-base estimation problem for the HRP-2 humanoid, assuming perfect whole-body kinematics but uncertain kinematics at the ankle, modeled as series elasticity concentrated at the foot. The authors represent this uncertainty as a pure rotation, assuming that linear translations of the compliance are negligible. Constraints on the system (such as contact points at the feet) are represented as virtual measurements as proposed by Xinjilefu and Atkeson [43]. The estimated state is composed of angular and linear position, velocity and acceleration of the floating base. The authors adopted a state transition model based on a simple integrator and a random walk for the floating base acceleration. The estimation strategy is based on an extended Kalman filter with measurements from the the robot gyroscope and accelerometer. Validation experiments consisted of

stabilizing the humanoid hand in an inertial reference frame, using the estimated floating-base position to compensate for the ankle elasticity; results show that the resulting positioning error reduces from $20[cm]$ to $2[cm]$ thanks to the proposed estimation strategy.

Rotella et al. [33] proposed a humanoid floating-base estimation strategy which extends the approach by Bloesch et al. [4] which was initially designed and tested on quadrupeds. The underlying idea is to include kinematic constraints (such as footholds) to infer the floating-base motions due to the (measured) joint motions. The proposed filter state contains the base position, the base velocity, the base orientation and the foot positions in the world reference frame. A quaternion based representation of the robot attitude allows one to obtain a global singularity-free representation. The system state is augmented to represent the accelerometer and gyroscope biases. An extended Kalman filter approach is proposed to solve the estimation problem. An interesting observability analysis is conducted by means of a taxonomy of rank deficiencies based on the different gait phases (single point foot, double foot point contact and flat foot). Reported localization errors in simulation are minimal: $0.05[m/s]$ on linear velocities, $0.05[rad]$ on orientations for a 120 second walking task.

Xinjilefu et al. [44, 45] proposed yet another floating-base estimation framework. In [44], authors consider simultaneously consider Problem 1 and 4, decomposing the global estimation problem in three sub-problems: base estimation, joint position estimation and joint velocity estimation. This approach is proven to simplify the computation burden of the full-body model estimation. The floating base estimation relies on a quaternion parametrization to avoid singularities. The joint velocities estimation includes both the contact constraints (31c) the system dynamics (31b), projected into the orthogonal complement of the constraints to cancel out contact forces as proposed in [22]. In [45], authors framed the same estimation problem as a quadratic programming, thus including inequality constraints in the estimation problem. Considered measurements include joint velocities, joint torques, contact forces, pelvis angular velocity and linear acceleration; as in previous approaches, kinematic constraints (such as footholds) are included as virtual measurements. Validation experiments conducted on the Atlas humanoid prove that the QP estimates demonstrate good tracking performance and less delay which allow to reduce the controller chattering.

5.5 Free-floating state estimation with range and vision sensors

The DARPA Robotics Challenge encouraged the competing teams to refine their systems (including state estimation) and to take them outside of the laboratory. A major challenge teams faced was the drift of the estimate of the robot's floating base (pelvis link) which in turn causes the position of objects of interest (e.g. doorways or stairway steps) to become inconsistent. Having tuned their performance, three teams competing with the Boston Dynamics Atlas robot [18, 46, 17] reported their open loop (i.e. using proprioceptive sensing only) drift rate to be approximately 1 cm per step (over distances of more 10 meters) at the end of the competition.

As described above approaches commonly use foothold kinematic constraints to estimate pelvis velocity. As a result, the performance of floating base estimation is dependent on the accuracy of contact classification and careful calibration and filtering of joint sensing. Drift is often characteristic — for example drifting forward slightly when walking forward backward.

To entirely avoid estimation drift it is common to explore approaches from the field of SLAM (Simultaneous Localization and Mapping) using range and vision sensors. The literature of the field is vast and we will not present it fully here, however an additional challenge for a robot such as a humanoid is that performance of many SLAM approaches is in fact correlated with the smoothness of the motion of the robot. Irregular and jerky motion usually contravenes assumptions underlying many approaches as well as causing motion blur in camera images and saturation in accelerometers.

Here we split approaches into two subcategories: those using LIDAR sensors and those using cameras.

5.5.1 LIDAR-based approaches

Laser Range Finders or LIDAR — because — of its long ranging performance and precise measurement capability is perhaps a more favorable approach to achieve drift-free localization over short ranges (less than 30 meters). [16] was a first approach achieving localization in a 2 floor setup with a 2D laser on the head of a NAO robot. The approach combined LIDAR Monte Carlo integration with information about the robot's height and orientation to produce an estimate in 6D.

While not carrying out localization, per say, the work of Nishiwaki et al. [24] is impressive. The authors use LIDAR to build a terrain map and then localize in it while continuously walking. This demonstration makes it clear

that the underlying proprioceptive odometry system has low drift.

For a full-size humanoid moving continuously walking, the approaches of Fallon et al. [11] and Koolen et al. [17] are similar. Both demonstrated continuous localization over 3D walking courses which require precise accuracy. Fallon et al. [11] used a Gaussian particle filter to estimate at 40Hz the position of the robot using each individual LIDAR scan for low latency. The Gaussian particle filter approach is interesting in that the filter sampled only in the states which are observable by the filter (for example the yaw of the robot, but not its pitch and roll). Meanwhile Koolen et al. [17] built upon the common iterative closest point (ICP) algorithm: first collating a point cloud using a tuned proprioceptive odometry algorithm and then aligning that point cloud to a prior map.

5.5.2 Camera-based approaches

Localization Because of the computational challenges of running real-time visual SLAM on-board a humanoid robot, initially researchers studied localization against prior maps. Early approaches such as [38] used stereo point clouds to localize, in this case, against height maps in 2D.

The work of Gonzalez-Aguirre et al. [15] explores active localization where the robot carried out gazing motions to actively improve the performance of a particle filter localizer. The approach used an inverse graphics method to render edges of a CAD model of a kitchen environment to match to edges detected by a head camera pointing at the real kitchen. The approach does not use point features and is to a degree engineered for the environment.

Much later Alcantarilla et al. [1] demonstrated monocular visual localization using a single camera in a more general manner. The approach localized against a precise prior map of visual point features with visibility of the 3D points in different poses calculated so as to speed up on-line feature matching. In common with many mapping systems for humanoids, the demonstrated motion was unfortunately in small 2m space but was interesting in that different lighting conditions and datasets were examined as was substantial comparison against PTAM, a commonly used visual SLAM system.

Finally, the approach of [5] is interesting in that it achieves localization without using either a camera or a LIDAR. The approach (demonstrated on a quadruped) used information from contact sensors to localize a quadruped on a detailed height-map using particle filtering.

Simultaneous Localization and Mapping With regard to the full problem of SLAM, the work of Stasse et al. [36] has a clear breakthrough — achieving the first SLAM system capable of carrying out loop closing in real-time running on-board a humanoid. The approach uses the seminal monoSLAM system to carry out map construction.

Yoon et al. [48] later developed an EKF-based SLAM with a stereo camera which could produce minimal drift for motion in a 4m box. This work clearly describes the development of robust visual descriptors as well as motion model switching to compensate for the uncertain motion of the humanoid. [27] is interesting as it demonstrates that even the monocular cameras on a small NAO robot are now able to use visual navigation. The approach used the camera input for visual odometry and in doing so is able to demonstrate a reduced rate of drift (rather than attempting to build a map).

Finally the work of Fallon et al. [12] is the first to use the output of a dense visual mapping system to produce terrain information which is used directly as input to the robot's control system for a continuously walking humanoid over rough terrain. Running a visual mapping system like this in real-time is a clear integration challenge.

5.6 Combined estimation

Some recent approaches to the estimation problem proposed a combined strategy, based on the idea of mixing different estimations simultaneously. Wu et al. [42] considered estimation and calibration simultaneously. Lowrey et al. [19] proposed an estimation strategy which incorporate a physics engine into the estimator to adjust the state estimates so that the inferred contact interactions are consistent with the observed accelerations. Nori et al. [25] proposed two alternative solutions to the simultaneous estimation of forces and joint related quantities (position, velocity and acceleration estimation). Hereafter we discuss this approach in details.

Wu et al. [42] developed a novel alternating optimization approach for simultaneously tracking space-time joint angles and calibrating associated parameters. Considered sensors are potentiometers at joints, motion capture markers attached to links and extension sensors on tendons. Simultaneously estimated variables are joint angles and sensor parameters. The latter include: (1) potentiometer gain and bias, (2) motion capture markers positions and (3) rotation/translation of the motion-capture reference frame with respect to the robot reference frame. The mixed estimation

strategy is similar to an EM-algorithm: a p-phase that estimates sensor parameters while keeping joint angle fixed and a q-phase that estimates joint angles using the updated sensor parameters. Comparing to the manual calibration done by manufacturer, the proposed estimation strategy reduced the error by half in all sensor types.

Lowrey et al. [19] considered accurate approach to state estimation which re-estimates the trajectory over a time window into the past and uses a recursive prior obtained from the previous time step via simulation with accurate physic engine (MuJoCo). The estimation is framed in a Bayesian-like framework (i.e. EKF) enhanced by updating the system’s history over a fixed number of steps, thus obtaining an increased accuracy and maintaining real-time performances. Exploited sensors include encoders, contact forces, 3D markers, and inertial measurement units. The estimation is conducted with a maximum-a-posteriori. The maximized function is the sum of the data likelihood and a corrective term which penalizes physical inconsistent estimations.

Nori et al. [25] proposed an alternative estimation framework which performs a multimodal estimation and takes into account the physical consistency of the computed quantities. Estimated variables include the system state (Problem 1: joint positions and velocities) and several dynamic quantities (Problem 2: joint torques, internal force/torques; Problem 3: external contact force/torques). Exploiting a Bayesian formulation, physical consistency is obtained by defining a suitable prior on estimated quantities. The estimation relies on the computation of the posterior probability by conditioning the prior on measurements from heterogeneous sensors: joint encoders, link gyroscopes and accelerometers and whole-body distributed 6-axis force/torque sensors. Experimental validation in simulation and on a real robot (the iCub humanoid) shows an accuracy of $0.51[deg/s]$ on joint velocities and $0.02[rad]$ on joint positions which drops down to $\approx 0.1[rad]$ on the real robot.

6 Future directions and open problems

Improving estimation reliability and accuracy, and in turn to provide the controller with improved input signals, will be a key requirement to increase the performance of the next generation of humanoids. Several improvements to the aforementioned approaches will be necessary. In this section we focus on future directions and open problems and their connections with previous literature.

Computational efficiency will be a core requirement for all future estimation strategies. Simple estimation problems like inverse and forward dynamics, have been extensively optimized from the computational point of view [13]. Significant simplifications have been achieved by exploiting the sparsity induced by the tree-like structure of humanoids. Similar advantages can be ported to a wider class of estimation problems as recently proposed in [25, Section V]. Other simplifications can be obtained by the decomposing the original non-convex estimation problem into simpler convex subproblems as proposed in [21] and [44].

Humanoids are free-floating constrained system and therefore constrained estimation will be another core requirement. The state of a constrained mechanical system evolves on a manifold whose topology is non-trivial, especially when in presence of contacts at multiple locations. The state estimation problem should take these constraints into account similarly to the approach proposed in Xinjilefu and Atkeson [43]. Extensions of this approach should be capable of accurate estimation even when constraints are broken or established. In general, there is a need for solutions to the problem of state estimation for hybrid (or switching) systems with applications to constrained mechanical systems as the one in equation (31).

Multi-modality and sensor fusion will play an important role in forthcoming applications. To provide a robust estimation infrastructure, existing techniques should be consolidated in a unique framework for sensor fusion of the different sensor modalities including inertial sensing, joint encoders, force and tactile sensing. Preliminary investigation of this have been already performed in [19] and in [25]. However, these methods have been tested mainly in simulation or in simple experimental settings, while in future applications the effort will be focused to bring these techniques in a real world scenario, and eventually improve them to cope with sensor specific issues.

Adaptive calibration and hyper-parameters tuning will be fundamental contributions to improve the estimation accuracy. It is often difficult to have an exact model of the controlled system. Adaptive calibration and tuning might reduce the different model uncertainties affecting the system (e.g. errors in the manufacturing process and unmodeled elements such as wires and cables). Current available approaches deal with sensor bias [37], sensor positions [32] and dynamic parameters [40]. Forthcoming applications based on multimodal sensor fusion will require dedicated techniques for estimating the relative accuracy of the different sensor modalities to weight their contribution in the estimation.

Improvements in the estimation can be obtained also by exploiting context dependent information, provided by

tactile sensing, RGB-D sensing and, in an indirect way, by the available world model. For example, if there is an external contact on a part of the robot not covered by the skin, the approximate location of the contact can be detected by the RGB-D sensor. It is worth noting that, with the notable exception of [20], RGB-D sensing has been traditionally limited to SLAM applications [11]. Similarly, combining tactile and proprioceptive sensing (i.e. joint encoders) can lead to an “embodied” representation of the surrounding environment as in the paper by Petrovskaya et al. [30].

Last but not least, the proposed estimation approaches needs to be validated both from the theoretical and practical point of view. At the theoretical level, observability analysis should be conducted. In general approaching a configuration where a given state is non-observable (i.e. the observability matrix loses rank) causes significant degradation in the estimation accuracy (i.e. the observability matrix has a relatively high condition number). Thus non-observable conditions should be avoided and estimation accuracy should be carefully analyzed. From a practical point of view, novel estimation strategies should to be compared with previous approaches to validate their improved accuracy. In practice this validation might be sometimes nontrivial. Bloesch et al. [4] proposed the use of an additional, more accurate, sensor as a validation ground-truth. [32] preferred an indirect validation based on showing that the estimation could improve the control performances (i.e. tracking error).

7 Conclusions

In this chapter we discussed previous sensor fusion approaches relevant to humanoid robotics. Focus is primarily on state estimation but includes other relevant problems: joint torques estimation, contact force estimation, contact position estimation and center of mass estimation. Previous approaches are discussed in relation to features which are relevant to humanoid applications: computational complexity, observability analysis, modeling accuracy and floating base parameterization. The paper concludes by observing that despite possible some open problems remain, the path paved by previous approaches indicate promising future direction which might overcome the limitations of current approaches.

8 Acknowledgment

This paper was supported by the FP7 EU projects CoDyCo (No. 600716 ICT 2011.2.1 Cognitive Systems and Robotics), and Koroibot (No. 611909 ICT-2013.2.1 Cognitive Systems and Robotics).

References

- [1] Pablo F. Alcantarilla, Olivier Stasse, Sebastien Druon, Luis M. Bergasa, and Frank Dellaert. How to localize humanoids with a single camera? *Autonomous Robots*, 34(1):47–71, 2013. ISSN 1573-7527.
- [2] Y. Aoustin, F. Plestan, and V. Lebastard. Experimental comparison of several posture estimation solutions for biped robot rabbit. In *Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on*, pages 1270–1275, May 2008. doi: 10.1109/ROBOT.2008.4543378.
- [3] M. Benallegue and F. Lamiraux. Humanoid flexibility deformation can be efficiently estimated using only inertial measurement units and contact information. In *2014 IEEE-RAS International Conference on Humanoid Robots*, pages 246–251, Nov 2014. doi: 10.1109/HUMANOIDS.2014.7041367.
- [4] Michael Bloesch, Marco Hutter, MH Hoepflinger, C David Remy, Christian Gehring, and R Siegwart. State Estimation for Legged Robots - Consistent Fusion of Leg Kinematics and IMU. In *Robotics: Science and Systems VIII*, 2012.
- [5] S. Chitta, P. Vemaza, R. Geykhman, and D. D. Lee. Proprioceptive localization for a quadrupedal robot on known terrain. In *Proceedings 2007 IEEE International Conference on Robotics and Automation*, pages 4582–4587, April 2007. doi: 10.1109/ROBOT.2007.364185.
- [6] C. Ciliberto, L. Fiorio, M. Maggiali, L. Natale, L. Rosasco, G. Metta, G. Sandini, and F. Nori. Exploiting global force torque measurements for local compliance estimation in tactile arrays. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3994–3999, Sept 2014. doi: 10.1109/IROS.2014.6943124.

- [7] A. De Luca, A. Albu-Schaffer, S. Haddadin, and G. Hirzinger. Collision detection and safe reaction with the dlr-iii lightweight manipulator arm. In *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1623–1630, Oct 2006. doi: 10.1109/IROS.2006.282053.
- [8] Andrea Del Prete, Simone Denei, Lorenzo Natale, Fulvio Mastrogiovanni, Francesco Nori, Giorgio Cannata, and Giorgio Metta. Skin spatial calibration using force/torque measurements. In *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3694–3700. IEEE, 2011.
- [9] Andrea Del Prete, Lorenzo Natale, Francesco Nori, and Giorgio Metta. Contact force estimations using tactile sensors and force/torque sensors. *Human Robot Interaction*, 2012:0–2, 2012.
- [10] Andrea Del Prete, Nicolas Mansard, Oscar E Ramos, Olivier Stasse, and Francesco Nori. Implementing torque control with high-ratio gear boxes and without joint-torque sensors. *International Journal of Humanoid Robotics*, 13(01):1550044, 2016.
- [11] Maurice Fallon, Matt Antone, Nicholas Roy, and Seth Teller. Drift-free humanoid state estimation fusing kinematic, inertial and lidar sensing. In *2014 IEEE-RAS International Conference on Humanoid Robots*, pages 112–119, Nov 2014. doi: 10.1109/HUMANOIDS.2014.7041346.
- [12] Maurice F. Fallon, Pat Marion, Robin Deits, Thomas Whelan, Matthew Antone, John McDonald, and Russ Tedrake. Continuous humanoid locomotion over uneven terrain using stereo fusion. In *2015 15th IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, Seoul, Korea, November 2015.
- [13] Roy Featherstone. *Rigid Body Dynamics Algorithms*. Springer, 2008.
- [14] Matteo Fumagalli, Serena Ivaldi, Marco Randazzo, Lorenzo Natale, Giorgio Metta, Giulio Sandini, and Francesco Nori. Force feedback exploiting tactile and proximal force/torque sensing. Theory and implementation on the humanoid robot iCub. *Autonomous Robots*, 33(4):381–398, 2012.
- [15] D. Gonzalez-Aguirre, M. Vollert, T. Asfour, and R. Dillmann. Robust real-time 6d active visual localization for humanoid robots. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2785–2791, May 2014. doi: 10.1109/ICRA.2014.6907258.
- [16] Armin Hornung, Kai M. Wurm, and Maren Bennewitz. Humanoid robot localization in complex indoor environments. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Taipei, Taiwan, October 2010.
- [17] Twan Koolen, Sylvain Bertrand, Gray Thomas, Tomas de Boer, Tingfan Wu, Jesper Smith, Johannes Engelsberger, and Jerry Pratt. Design of a momentum-based control framework and application to the humanoid robot atlas. *International Journal of Humanoid Robotics*, 13, 2016.
- [18] Scott Kuindersma, Robin Deits, Maurice Fallon, Andres Valenzuela, Hongkai Dai, Frank Permenter, Twan Koolen, Pat Marion, and Russ Tedrake. Optimization-based locomotion planning, estimation, and control design for atlas. *Autonomous Robots*, 40:429–455, 2016.
- [19] K. Lowrey, S. Kolev, Y. Tassa, T. Erez, and E. Todorov. Physically-consistent sensor fusion in contact-rich behaviors. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1656–1662, Sept 2014. doi: 10.1109/IROS.2014.6942777.
- [20] E. Magrini, F. Flacco, and A. De Luca. Estimation of contact forces using a virtual force sensor. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2126–2133, Sept 2014. doi: 10.1109/IROS.2014.6942848.
- [21] Lucas Manuelli and Russ Tedrake. Localizing external contact using proprioceptive sensors: The contact particle filter. *Under review*, 2016.
- [22] M. Mistry, J. Buchli, and S. Schaal. Inverse dynamics control of floating base systems using orthogonal decomposition. In *Robotics and Automation (ICRA), 2010 IEEE International Conference on*, pages 3406–3412, May 2010. doi: 10.1109/ROBOT.2010.5509646.

- [23] Richard M Murray, Zexiang Li, S Shankar Sastry, and S Shankara Sastry. *A mathematical introduction to robotic manipulation*. CRC press, 1994.
- [24] Koichi Nishiwaki, Joel Chestnutt, and Satoshi Kagami. Autonomous navigation of a humanoid robot over unknown rough terrain using a laser range sensor. 31:1251–1262, 2012.
- [25] F. Nori, N. Kuppuswamy, and S. Traversaro. Simultaneous state and dynamics estimation in articulated structures. In *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*, pages 3380–3386, Sept 2015. doi: 10.1109/IROS.2015.7353848.
- [26] Francesco Nori, Silvio Traversaro, Jorhabib Eljaik, Francesco Romano, Andrea Del Prete, and Daniele Pucci. icub whole-body control through force regulation on rigid noncoplanar contacts. *Frontiers in Robotics and AI*, 2(6), 2015. ISSN 2296-9144. doi: 10.3389/frobt.2015.00006. URL http://www.frontiersin.org/humanoid_robotics/10.3389/frobt.2015.00006/abstract.
- [27] Giuseppe Oriolo, Antonio Paolillo, Lorenzo Rosa, and Marilena Vendittelli. Humanoid odometric localization integrating kinematic, inertial and visual information. *Autonomous Robots*, 40(5):867–879, 2016. ISSN 1573-7527. doi: 10.1007/s10514-015-9498-0. URL <http://dx.doi.org/10.1007/s10514-015-9498-0>.
- [28] Jaeheung Park and O. Khatib. Multi-link multi-contact force control for manipulators. In *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*, pages 3613–3618, April 2005. doi: 10.1109/ROBOT.2005.1570670.
- [29] Y. J. Park and W. K. Chung. External torque sensing algorithm for flexible-joint robot based on disturbance observer structure. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 4735–4741, Sept 2014. doi: 10.1109/IROS.2014.6943236.
- [30] A. Petrovskaya, J. Park, and O. Khatib. Probabilistic estimation of whole body contacts for multi-contact robot control. In *Proceedings 2007 IEEE International Conference on Robotics and Automation*, pages 568–573, April 2007. doi: 10.1109/ROBOT.2007.363047.
- [31] Lekskulchai Pongsak, Masafumi Okada, and Yoshihiko Nakamura. Optimal filtering for humanoid robot state estimators. In *Proceedings of SICE System Integration Division Annual Conference (SI2002)*, 2002.
- [32] N. Rotella, S. Mason, S. Schaal, and L. Righetti. Inertial sensor-based humanoid joint state estimation. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1825–1831, May 2016. doi: 10.1109/ICRA.2016.7487328.
- [33] Nicholas Rotella, Michael Bloesch, Ludovic Righetti, and Stefan Schaal. State estimation for a humanoid robot. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 952–958. IEEE, 2014.
- [34] B Siciliano and L Sciacivco. *Robotics: modelling, planning and control*. Springer, 2009.
- [35] V. Sotoudehnejad, A. Takhmar, M. R. Kermani, and I. G. Polushin. Counteracting modeling errors for sensitive observer-based manipulator collision detection. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 4315–4320, Oct 2012. doi: 10.1109/IROS.2012.6386198.
- [36] Olivier Stasse, Andrew J. Davison, Ramzi Sellaouti, and Kazuhito Yokoi. Real-time 3D SLAM for a humanoid robot considering pattern generator information. In *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2006.
- [37] Benjamin J. Stephens. State estimation for force-controlled humanoid balance using simple models in the presence of modeling error. In *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, pages 3994–3999, 2011.
- [38] S. Thompson and S. Kagami. Humanoid robot localisation using stereo vision. In *5th IEEE-RAS International Conference on Humanoid Robots, 2005.*, pages 19–25, Dec 2005. doi: 10.1109/ICHR.2005.1573539.

- [39] Luc Le Tien and Alin Albu-Schäffer. Friction Observer and Compensation for Control of Robots with Joint Torque Measurement. In *Intelligent Robots and Systems, IROS. IEEE/RSJ International Conference on*, pages 22–26, 2008.
- [40] Silvio Traversaro, Andrea Del Prete, Serena Ivaldi, and Francesco Nori. Inertial parameters identification and joint torques estimation with proximal force/torque sensing. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2105–2110. IEEE, 2015.
- [41] Juho Vihonen, Janne Honkakorpi, Jouni Mattila, and Ari Visa. Geometry-aided angular acceleration sensing of rigid multi-body manipulator using mems rate gyros and linear accelerometers. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2013.
- [42] Tingfan Wu, Y. Tassa, V. Kumar, J. Movellan, and E. Todorov. Stac: Simultaneous tracking and calibration. In *2013 13th IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, pages 469–476, Oct 2013. doi: 10.1109/HUMANOIDS.2013.7030016.
- [43] Ben Xinjilefu and Christopher G. Atkeson. State estimation of a walking humanoid robot. In *IROS*, pages 3693–3699. IEEE, 2012.
- [44] Ben Xinjilefu, Siyuan Feng, Weiwei Huang, and Christopher Atkeson. Decoupled state estimation for humanoids using full-body dynamics. In *IEEE/RAS 2014 IEEE International Conference on Robotics and Automation (ICRA 2014)*, 2014.
- [45] Ben Xinjilefu, Feng Siyuan, and Christopher G. Atkeson. Dynamic state estimation using quadratic programming. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 989–994. IEEE, 2014.
- [46] Ben Xinjilefu, Siyuan Feng, and Christopher Atkeson. Center of mass estimator for humanoids and its application in modelling error compensation, fall detection and prevention. In *International Conference on Humanoid Robotics*, Seoul, Korea, November 2015.
- [47] Ben Xinjilefu, Siyuan Feng, and Christopher G. Atkeson. A distributed mems gyro network for joint velocity estimation. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1879–1884, May 2016. doi: 10.1109/ICRA.2016.7487334.
- [48] Sukjune Yoon, Seungyong Hyung, Minhyung Lee, Kyung Shik Roh, SungHwan Ahn, Andrew Gee, Pished Bunnun, Andrew Calway, and Waterio W. Mayol-Cuevas. Real-time 3d simultaneous localization and map-building for a dynamic walking humanoid robot. *Advanced Robotics*, 27(10):759–772, 2013.
- [49] H. Zhang, S. Ahmad, and G. Liu. Torque estimation for robotic joint with harmonic drive transmission based on position measurements. *IEEE Transactions on Robotics*, 31(2):322–330, April 2015. ISSN 1552-3098. doi: 10.1109/TRO.2015.2402511.