A Tactile Sensor-Based EKF Estimator to Predict Object-Contact States During Whole-Body MultiContact Manipulation

Rajesh Subburaman[1], Francesco Rachiglia[2], and Olivier Stasse[1,3]

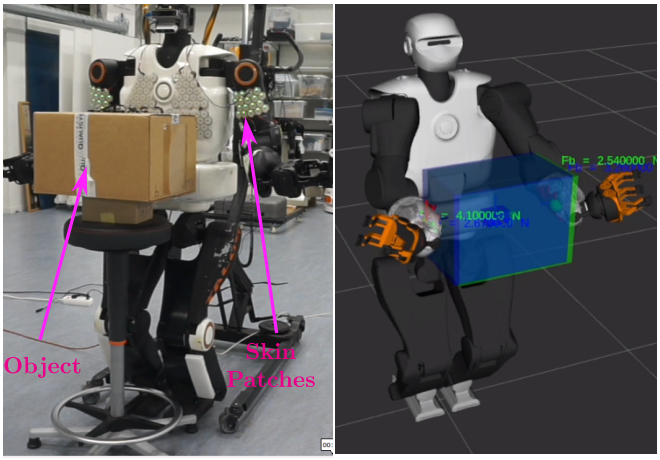
***Abstract*— We present an extended Kalman filter–based framework for online estimation of object and contact states during whole-body multi-contact manipulation of large objects with humanoids. By combining tactile skin patches, joint encoders, forward kinematics, and an IMU, the estimator tracks object and contact pose, twist, and forces in real time. Unlike conventional end-effector-centric pose estimation approaches, our approach leverages whole-body contact data available from skin patches to estimate the states of large and heavy objects during dynamic manipulation. Experiments on the TALOS humanoid performing dynamic grasp and lift tasks demonstrate promising results on tracking of object-contact states, with potential to enhance manipulation stability, and support complex whole-body manipulations including contact shuffling and can further broaden humanoid applications.**

I. Introduction

Humanoids are designed to operate in human environments and handle objects and tools made for humans. While robotic locomotion has achieved robustness over complex terrains, manipulation still lags behind human capabilities. Most existing approaches are end-effector–centric [1], [2] and struggle with objects of varying geometry and dynamics. In contrast, whole-body contact [3], [4] can improve efficiency, reduce joint loads, and enable handling of large, heavy objects, yet it remains underexplored—partly due to challenges such as reliable object state estimation relative to the robot. This information is especially critical for whole-body manipulation, affecting performance, stability, and tasks like dynamic contact switching and parameter estimation [5]. Recent advances in lightweight tactile sensing [6] enable whole-body contact feedback. Leveraging this, we propose an extended Kalman filter (EKF)-based framework that fuses tactile data with joint encoders, forward kinematics, and IMU measurements to estimate object and contact pose, twist, and forces online during dynamic manipulation. Prior works [7]–[9] has largely focused on estimating the pose of small handheld objects, with limited dynamic scenarios [10]. Our framework addresses these gaps, and preliminary results demonstrate robust performance.

II. Methodology

The estimation approach relies on four assumptions: (i) contacts are modeled as soft point contacts, with three force components and one moment along the contact normal; (ii) the estimator is initialized only after all intended skin patches are in contact with the object (grasp phase); (iii) the object's initial pose is known at the grasp phase and used for initialization; and (iv) the object reference frame is placed at its center of mass, coinciding with its geometric center.



(a) (b)

Fig. 1: (a) The carton box manipulation experiment setup with the TALOS humanoid. (b) The proposed EKF estimator and the ground truth data are visualized in rviz.

1) System Description: The experimental setup uses the TALOS humanoid [11], a 32-DoF robot (95 kg, 1.75 m tall), shown in Fig. 1. The manipulated object is a carton box (0.44 × 0.35 × 0.27 m, 0.75 kg). Whole-body multi-contact manipulation is performed using the torque-based controller from [5]. TALOS is equipped with joint torque sensors and encoders, as well as force-torque sensors at the wrists and ankles. It also features 12 upper-body skin patches composed of hexagonal cells, each integrating accelerometers, temperature, proximity, and capacitance-based pressure sensors, and providing pose information relative to their respective root. In this work, only the two forearm patches are used.

2) State Representation: In [5], WBMC relies on contact references derived from the object's desired pose and twist, making accurate estimation of both object and contact states essential. Additionally, contact force estimation is required to maintain the desired grasp. Accordingly, the estimated state $\mathbf{x}$ includes the following:

$$\mathbf{x} = \begin{bmatrix} \mathbf{r}_c^i & \mathbf{v}_c^i & \mathbf{q}_c^i & \mathbf{r}_o & \mathbf{v}_o & \mathbf{q}_o & \omega_o & \mathbf{f}_c^i \end{bmatrix}, \quad (1)$$

where $\mathbf{r}_c^i$, $\mathbf{v}_c^i$, and $\mathbf{q}_c^i$ denote the position, velocity, and orientation of the i^{th} contact, while $\mathbf{r}_o$, $\mathbf{v}_o$, and $\mathbf{q}_o$ represent those of the object. All states are expressed in the world frame w, except the contact forces $\mathbf{f}_c^i$ that are defined in the contact frame. Rotation are represented using quaternions.

3) Process Model: In continuous time, the nonlinear prediction model for object–contact state evolution is derived

[1] Gepetto, LAAS-CNRS, Universite de Toulouse, France
[2] Ecole Centrale de Nantes, France
[3] Artifical and Natural Intelligence Toulouse Institute, France

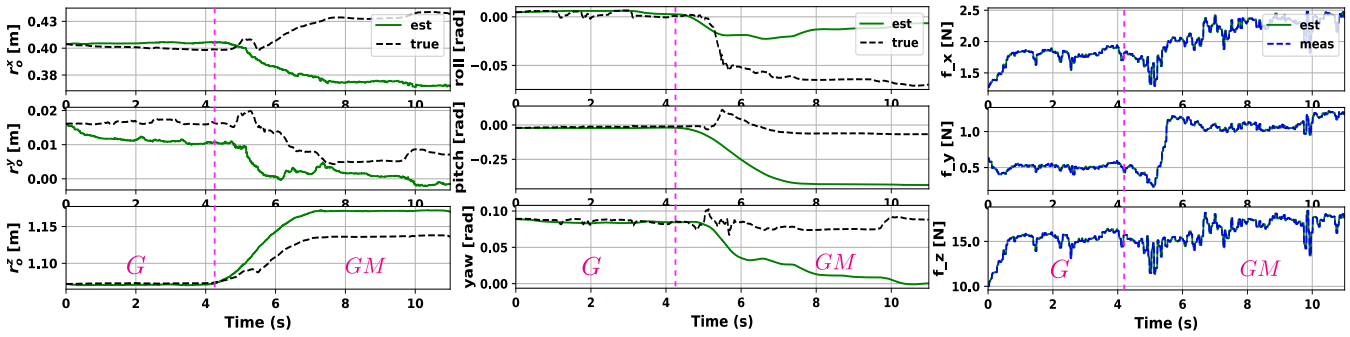


Fig. 2: The proposed EKF estimation results of $\mathbf{r}_{obj}$, $\mathbf{q}_{obj}$ (in euler), and $\mathbf{f}_c$ (left forearm contact) are compared here with the mocap data (true).

using three-axis accelerometer data from each skin cell and the grasp matrix $\mathbf{G}$, which relates ${}^b\nu_c = \mathbf{G}^T {}^b\nu_o$. Here, ${}^b\nu_o \in \mathbb{R}^6$ and ${}^b\nu_c \in \mathbb{R}^{4\times n_c}$ denote the object and contact twists in the body frame b. For more details, please refer [12]. The resulting prediction model follows as.

$$
\begin{aligned}
\dot{\mathbf{r}}_c^i &= \mathbf{v}_c^i, \\
\dot{\mathbf{v}}_c^i &= \mathbf{a}_c^i = {}^c\mathbf{C}_{\mathrm{w}}^T \tilde{\mathbf{e}} + \mathbf{g} + \mathbf{w}_a, \\
\dot{\mathbf{q}}_c^i &= 0.5 \begin{bmatrix} \omega_c^i & 0 \end{bmatrix}^T \otimes \mathbf{q}_c^i, \\
\dot{\mathbf{r}}_o &= {}^{\mathrm{o}}\mathbf{C}_{\mathrm{w}}^T (\mathbf{G}\mathbf{v}_c) + \mathbf{w}_o, \\
\dot{\mathbf{v}}_o &= {}^{\mathrm{o}}\mathbf{C}_{\mathrm{w}}^T (\mathbf{G}\mathbf{a}_c) + \mathbf{w}_a, \\
\dot{\mathbf{q}}_o &= 0.5 \begin{bmatrix} \omega_o & 0 \end{bmatrix}^T \otimes \mathbf{q}_o, \\
\dot{\omega}_o &= \mathbf{w}_{\omega_o}, \\
\dot{\mathbf{f}}_c^i &= \mathbf{w}_f,
\end{aligned}
\tag{2}
$$

where $\mathbf{w}_\star$, $\tilde{\mathbf{e}}$, $\mathbf{g}$, and $\mathbf{C}$ denote process noise, resultant acceleration of the skin patch, gravity, and the rotation matrix with respect to the world frame, respectively. Due to the absence of a gyroscope in each skin cell and the difficulty in estimating contact forces, $\dot{\omega}_o$ and $\dot{\mathbf{f}}_c^i$ are taken as random noises. The object kinematics is inferred from the contact kinematics via $\mathbf{G}$, as in [12].

4) Measurement Model: The nonlinear measurement model uses data from joint encoders, skin patches, the waist-IMU–based base state estimator, and the robot's forward kinematics. It relates object–contact states to sensor readings and includes additive Gaussian noise to account for sensor inaccuracies. The model thus obtained is given below.

$$
\begin{aligned}
\mathbf{r}_c^i &= {}^w\mathbf{T}_l \, {}^l\mathbf{r}_c^i, \\
\mathbf{v}_c^i &= \mathbf{Ad}_{\mathrm{T}_c^w} \mathbf{Ad}_{\mathrm{T}_l^c} (\mathbf{J}_l \dot{\theta}) + \mathbf{n}_{v_c}, \\
\mathbf{q}_c^i &= {}^w\mathbf{q}_l \, {}^l\mathbf{q}_c + \mathbf{n}_{q_c}, \\
\mathbf{r}_o &= {}^w\mathbf{C}_c \, {}^c\mathbf{d}_o^i + \mathbf{r}_c^i + \mathbf{n}_o, \\
\mathbf{v}_o &= \mathbf{G}\mathbf{v}_c + \mathbf{n}_{v_o}, \\
\mathbf{q}_o &= {}^w\mathbf{q}_c ({}^o\mathbf{q}_c)^{-1} + \mathbf{n}_{q_o}, \\
\omega_o &= {}^w\mathbf{C}_l (\mathbf{J}_l^a \dot{\theta}) + \mathbf{n}_{\omega_o}, \\
\mathbf{f}_c^i &= {}^c\mathbf{C}_l \sum_{k=1}^{n_{sc}} ({}^l\mathbf{d}_{sc} \times {}^l\mathbf{f}_{sc,z}) + \mathbf{n}_{f_c},
\end{aligned}
\tag{3}
$$

where $\mathbf{n}_\star$ denotes measurement noise, ${}^a\mathbf{T}_b$ is the transformation matrix from b to a, and $\mathbf{J}_l$ is the Jacobian of the link containing the i^{th} patch and its angular and linear components are denoted by l and a, respectively. In $\mathbf{v}_c^i$, $\mathrm{Ad}_{\mathrm{T}_b^a}$ is the adjoint of $\mathbf{T}_b^a$. In the last equation, ${}^l\mathbf{d}_k$ and ${}^l\mathbf{f}_k$ are the position and normal force of the k^{th} skin cell in link frame l. Here, $\mathbf{f}_k = [0\ 0\ p_k \cdot K_p]^{\mathrm{T}}$, with p_k the sensor's normal pressure value and K_p the calibration constant.

III. Experiments and Results

The experimental setup considered for the evaluation of the proposed estimator is as shown in Fig.1(a). The whole-body motion involves three phases: approach, grasp, and grasp-move, with the estimator starting at the grasp phase. During grasp-move, the object is manipulated by [-0.025, 0, 0.1] from its initial pose, controlled via predefined contacts at the forearm patches. A quintic polynomial Cartesian trajectory generator is used for all phases in the WBMC controller. For details, see [5].

1) Results: The proposed estimator's tracking of $\mathbf{r}_o$, $\mathbf{q}_o$, and $\mathbf{f}_c^0$ (contact force on the left forearm patch) is shown in Fig. 2, with grasp and grasp-move phases denoted by G and GM. During G, $\mathbf{r}_o$ remains mostly static except for $\hat{r}_o^y$, which reflects object deformation captured by the estimator using kinematics and skin patch data, unlike the mocap $\bar{r}_o^y$ based on a rigid model. In GM, $\hat{r}_o^x$ and $\hat{r}_o^z$ follow their trajectories, differing from mocap due to rotational slip along the Y-axis; this slip appears in $\bar{\mathbf{r}}_o$ but not in $\hat{\mathbf{r}}_o$ since it occurs about the contacts. The rotational slip is also visible in ϕ_o^y, where $\hat{\phi}_o^y$ rotates negatively as the robot moves the object closer to its torso to maintain balance, causing elbow flexion. Please see the video given here:Experiment Video. For $\mathbf{f}_c^0$, estimates closely follow skin patch measurements since the process model treats them as Gaussian noise. With the contact frame Z-axis aligned along the grasp, the desired force is 15 N; f_z rises during G and stabilizes near 15 N throughout GM.

IV. Conclusion and Future Work

In this work, we have proposed a tactile skin patch-based EKF estimator to estimate the object-contact states during whole-body multi-contact manipulation of a large object. The preliminary results of the proposed estimator are presented here and discussed briefly. The results are found to be promising but at the same time raises several open questions such as handling the rotational slip, and inclusion of joint torque sensors for better contact force estimates. As future work, we intend to improve the whole-body multi-contact control performance to improve the estimation accuracy, incorporate joint-torque sensors in the estimator, and evaluate the proposed estimator during more dynamic motions.

REFERENCES

[1] K. Harada, S. Kajita, H. Saito, M. Morisawa, F. Kanehiro, K. Fujiwara, K. Kaneko, and H. Hirukawa, "A humanoid robot carrying a heavy object," in *Proceedings of the 2005 IEEE international conference on robotics and automation*. IEEE, 2005, pp. 1712–1717.

[2] S. Y. Shin and C. Kim, "Human-like motion generation and control for humanoid's dual arm object manipulation," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 4, pp. 2265–2276, 2014.

[3] S. Nozawa, R. Ueda, Y. Kakiuchi, K. Okada, and M. Inaba, "A full-body motion control method for a humanoid robot based on on-line estimation of the operational force of an object with an unknown weight," in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2010, pp. 2684–2691.

[4] M. Florek-Jasińska, T. Wimböck, and C. Ott, "Humanoid compliant whole arm dexterous manipulation: control design and experiments," in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2014, pp. 1616–1621.

[5] R. Subburaman and O. Stasse, "A whole-body multi contact large object manipulation and estimation framework for humanoids using skin patches," in *2025 IEEE-RAS 24th International Conference on Humanoid Robots (Humanoids)*. IEEE, 2025, pp. 1–8.

[6] G. Cheng, E. Dean-Leon, F. Bergner, J. R. G. Olvera, Q. Leboutet, and P. Mittendorfer, "A comprehensive realization of robot skin: Sensors, sensing, control, and applications," *Proceedings of the IEEE*, vol. 107, no. 10, pp. 2034–2051, 2019.

[7] P. Hebert, N. Hudson, J. Ma, and J. Burdick, "Fusion of stereo vision, force-torque, and joint sensors for estimation of in-hand object location," in *2011 IEEE International Conference on Robotics and Automation*. IEEE, 2011, pp. 5935–5941.

[8] L. Röstel, L. Sievers, J. Pitz, and B. Bäuml, "Learning a state estimator for tactile in-hand manipulation," in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2022, pp. 4749–4756.

[9] H. Zhang, Q. Wang, A. Song, J. Lai, D. Ma, C. Zhao, G. Gao, Q. Li, and J. Zhang, "Tacsem: Tactile sensing, 3-d state estimation and in-hand manipulation of a soft hand," *IEEE/ASME Transactions on Mechatronics*, 2025.

[10] P. Hebert, N. Hudson, J. Ma, and J. W. Burdick, "Dual arm estimation for coordinated bimanual manipulation," in *2013 IEEE International Conference on Robotics and Automation*. IEEE, 2013, pp. 120–125.

[11] O. Stasse, T. Flayols, R. Budhiraja, K. Giraud-Esclasse, J. Carpentier, J. Mirabel, A. Del Prete, P. Souères, N. Mansard, F. Lamiraux, *et al.*, "Talos: A new humanoid research platform targeted for industrial applications," in *2017 IEEE-RAS 17th International Conference on Humanoid Robotics (Humanoids)*. IEEE, 2017, pp. 689–695.

[12] R. M. Murray, Z. Li, and S. S. Sastry, *A mathematical introduction to robotic manipulation, Chapter 5*. CRC press, 2017.