Kinematics Correspondence As Inexact Graph Matching

An T. Le
Intelligent Autonomous Systems
Technical University of Darmstadt
`an@robot-learning.de`

Kay Pompetzki
Intelligent Autonomous Systems
Technical University of Darmstadt
`kay@robot-learning.de`

Jan Peters
Intelligent Autonomous Systems
Technical University of Darmstadt
Centre for Cognitive Science
German Research Center for AI (DFKI)
Hessian.AI
`jan@robot-learning.de`

Armin Biess
Q (Cue) Ltd.
`armin.biess@q.ai`

Abstract

Imitation is a major aspect of intelligence behavior, which provides a versatile and rapid mechanism to transfer motor skills from one intelligent agent (e.g., human, animal, or robot) to another. One major challenge in imitation learning is the correspondence problem: establishing corresponding states and actions between an expert demonstrating the task and a learner trying to imitate the task when their embodiments are dissimilar (morphology, dynamics, degrees of freedom, etc.). Many existing approaches circumvent the correspondence problem by directly providing demonstrations of the robot learner, for example, in kinesthetic teaching or teleoperation. These methods require robot-specific proprioception, which is not always available. In this study, we investigate the correspondence problem between dissimilar embodiments. In particular, we propose a correspondence divergence between embodiments and derive an imitation policy via the proposed divergence in the reactive motion generation or inverse reinforcement learning settings. We conducted the first set of experiments with increasing complexity of embodiments, showing that the approach is well suited for identifying morphology correspondence in robot imitation.
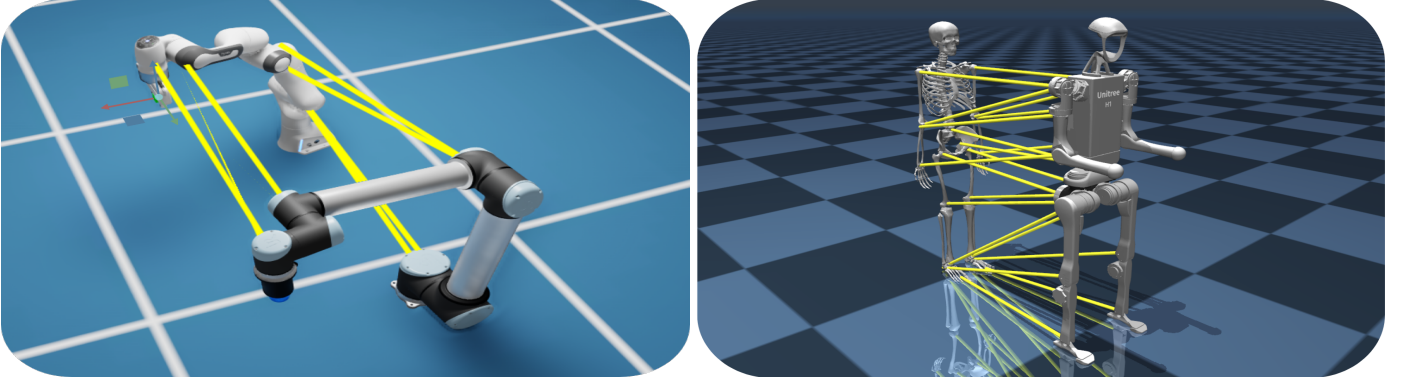
1 Introduction

Figure 1: The optimal correspondence $\boldsymbol{X}^*$ solving Eq. (1) represents the identified correspondences (yellow lines) between Panda & UR10e, and Skeleton & Unitree H1. More details can be found at `https://sites.google.com/view/igmcor`.

Imitation learning [1] has led to remarkable accomplishments, spanning from intricate helicopter acrobatics [2] and rapid arm dexterity [3] to the finesse of haptic control [4, 5], expressive gestures [6], precise manipulation [7–9], and even the complexity of legged locomotion [9, 10]. Imitation learning algorithms have been extensively studied and were recently summarised in a comprehensive overview [11]. Interestingly, despite the impressive advancements in developing novel robotic motor skills, fundamental challenges and open problems within imitation learning have persisted for decades. At the heart of these core inquiries lies the intricate *correspondence problem*: How can an agent (i.e., the learner) replicate a behavior observed in another agent (referred to as the expert)? This challenge gains complexity due to the disparate kinematic and dynamic constraints adhered to by the two agents. These constraints include body morphology, degrees of freedom (DOFs), potential self-collisions, joints, actuators, and torque limitations. In essence, the two embodiments exist within distinct state and action spaces [12], presenting a *morphology correspondence* that remains unresolved.

Recently, due to the scarcity of demonstration data for specific robot embodiments, data-driven motion retargeting has been used to circumvent manual mapping by state distribution matching [13–16] or latent space alignment [17–21]. Moreover, the trend of large robotics datasets [22–24] demands a scalable approach to identify similar contexts and skills, thereby supplying the data streams to robotics foundational models [25].

In this work, we investigate the correspondence problem in detail by representing the embodiment as a structured graph and deriving a similarity measure on these graphs based on some isometry heuristics. In particular, we propose a representation of embodiment as a geometrics graph with feature functions for graph comparison. Then, we formulate the correspondence as an inexact graph-matching problem to solve the correspondence assignments between dissimilar embodiments, forming a similarity divergence between embodiments. In the first set of experiments, we derive a correspondence Riemannian Motion Policy (RMP) for reactive imitation within the RMPflow [26], thus additionally incorporating robot-specific objectives, such as self-collision avoidance, joint-limit, or obstacle avoidance. In the inverse RL setting, we calibrate correspondence between the human skeleton and Unitree H1 as motion retargeting for walking tasks trained with Proximal Policy Optimization (PPO) [27].

2 Correspondence As Inexact Graph Matching

Here, we present the correspondence problem as an inexact graph matching problem [28], whose solution is the correspondence assignments between considered embodiments.

2.1 Embodiment Representation As Geometric Graph

Our choice of embodiment representation is motivated by imposing the following requirements on the model descriptions: (i) *generic* - the model can be applied to a variety of physical embodiments; (ii) *kinematics*- it can represent the shape and motion of the embodiment; (iii) *extrinsic* - embodiment features such as links or joint poses can be measured w.r.t. to an extrinsic reference frame; (iv) *metric-compatible* - it allows the introduction of a distance measure across dissimilar embodiments with different DoFs. These requirements are naturally satisfied by defining the embodiment graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, d^v, d^s, d^e)$, where the robot joint poses represent the node set $\mathcal{V} = \{\boldsymbol{T}_i\}_{i=1}^n$, $\boldsymbol{T} \in \mathrm{SE}(3)$, the robot physical links represent edge set $\mathcal{E} = \{e_k = (i, j) | 0 < i, j \leq n\}_{k=1}^m$, and the geometric functions. The global node feature function $d^v : \mathrm{SE}(3) \times \mathrm{SE}(3) \to \mathbb{R}$ is simply the distance metric defined on $\mathrm{SE}(3)$ manifold [29], representing node similarity as joint pose comparison w.r.t. a world frame. The local node feature function $d^s : \mathcal{V} \to \mathbb{N}$ defines local node heuristics measuring the physical embodiment topology such as the number of robot links connected to the ith-joint $d^s(i) = \mathrm{deg}(i)$

(i.e., the node degree). The specific edge feature function $d^e : \mathcal{E} \to \mathbb{R}$ computes physical link length, representing edge similarity as an embodiment structure comparison.

We construct the embodiment graph $\mathcal{G}$ by representing a node to each joint of the robot body. The edges are instantiated where there exists a physical link between the joints. For an embodiment with M-DoFs and generalized joint configuration $\boldsymbol{q} \in \mathbb{R}^M$, the joint poses w.r.t. a world frame $\mathcal{V}(\boldsymbol{q})$ is computed by a forward kinematic (FK) function mapping. Note that the initial joint poses can be freely chosen by the user or follow the robot description files. Embodiment graph equality, $\mathcal{G}_1 = \mathcal{G}_2$, follows if $|\mathcal{V}_1| = |\mathcal{V}_2|$, $d^v(\boldsymbol{T}_1^i, \boldsymbol{T}_2^i) = 0, d_1^s(i) = d_2^s(i) \, \forall i \in \{1, \ldots, n\}$ and $\mathcal{E}_1 = \mathcal{E}_2$, $d_1^e(k) = d_2^e(k)$, $\forall k \in \{1, \ldots, m\}$.

2.2 The Embodiment Correspondence As Graph Matching

We define the correspondence between two embodiment graphs $\mathcal{G}_1, \mathcal{G}_2$ as an injective function $c : \mathcal{V}_1 \to \mathcal{V}_2$, which maps corresponding nodes between embodiment graphs. We hypothesize that the correspondence problem in imitation learning can be formulated as an inexact Graph Matching (GM) problem [28], where a bijective map between embodiment nodes usually does not exist (e.g., different number of nodes between graphs $n_1 \neq n_2$). The inexact GM tries to find correspondence concerning minimum rather than zero-distortion, or, on the other hand, maximizing affinity between graphs.

First, the correspondence map c can be reformulated as the binary assignment matrix $\boldsymbol{X} \in \{0, 1\}^{n_1 \times n_2}$, and the graph topology of $\mathcal{G}$ can be represented by the node-edge incident matrix $\boldsymbol{G} \in \{0, 1\}^{n \times m}$, where $\boldsymbol{G}_{i,k} = \boldsymbol{G}_{j,k} = 1$ if the ith and j^{th}-nodes are connected by the kth-edge. The correspondence problem between $\mathcal{G}_1, \mathcal{G}_2$ can be formulated as Lawler's Quadratic Assignment Problem (QAP)

$$\max_{\boldsymbol{X}} \text{vec}(\boldsymbol{X})^\intercal \boldsymbol{K} \text{vec}(\boldsymbol{X}) \text{ s.t. } \boldsymbol{X} \in \{0, 1\}^{n_1 \times n_2}, \boldsymbol{X} \boldsymbol{1}_{n_2} = \boldsymbol{1}_{n_1}, \boldsymbol{X}^\intercal \boldsymbol{1}_{n_1} \leq \boldsymbol{1}_{n_2} \tag{1}$$

with $\boldsymbol{K} \in \mathbb{R}^{n_1 n_2 \times n_1 n_2}$ as the second-order affinity matrix, and $\text{vec}(\cdot)$ operator vectorizes matrices. The constraints enforce one-to-one node mapping between graphs, where a node in $\mathcal{G}_1$ can match at most one node in $\mathcal{G}_2$. Following [30], $\boldsymbol{K}$ can be factorized into

$$\boldsymbol{K} = (\boldsymbol{H}_2 \otimes \boldsymbol{H}_1) \text{diag}(\text{vec}(\boldsymbol{L}))(\boldsymbol{H}_2 \otimes \boldsymbol{H}_1)^\intercal,$$
$$\text{s.t. } \boldsymbol{H}_1 = [\boldsymbol{G}_1, \boldsymbol{I}_{n_1}] \in \{0, 1\}^{n_1 \times (m_1 + n_1)}, \boldsymbol{H}_2 = [\boldsymbol{G}_2, \boldsymbol{I}_{n_2}] \in \{0, 1\}^{n_2 \times (m_2 + n_2)},$$
$$\boldsymbol{L} = \begin{bmatrix} \boldsymbol{K}^e & -\boldsymbol{K}^e \boldsymbol{G}_2^\intercal \\ -\boldsymbol{G}_1 \boldsymbol{K}^e & \boldsymbol{G}_1 \boldsymbol{K}^e \boldsymbol{G}_2^\intercal + \boldsymbol{K}^v \end{bmatrix} \in \mathbb{R}^{(m_1 + n_1) \times (m_2 + n_2)}, \boldsymbol{I}_n \in \mathbb{R}^{n \times n} \text{ is identity matrix.} \tag{2}$$

The embodiment graph features can be introduced in the factorized problem Eq. (2) via the edge-affinity and node-affinity matrices $\boldsymbol{K}^e \in \mathbb{R}^{m_1 \times m_2}, \boldsymbol{K}^v \in \mathbb{R}^{n_1 \times n_2}$ with elements as Gaussian affinity score computed from the feature functions

$$\boldsymbol{K}_{i,j}^v = \exp\left(-\frac{w_s(d_1^s(i) - d_2^s(j))^2 + w_v d^v(\boldsymbol{T}_1^i, \boldsymbol{T}_2^j)^2}{\sigma^v}\right), i \in \{1, \ldots, n_1\}, j \in \{1, \ldots, n_2\}$$
$$\boldsymbol{K}_{i,j}^e = \exp\left(-\frac{(d_1^e(k) - d_2^e(l))^2}{\sigma^e}\right), k \in \{1, \ldots, m_1\}, l \in \{1, \ldots, m_2\} \tag{3}$$

where $\sigma^e, \sigma^v > 0$ are hyperparameters for node/edge affinity contributions, and $w_v, w_s > 0$ are hyperparameters for global/local node feature contributions.

There are several solvers for the objective Eq. (2) in the graph-matching literature [28, 31]. However, the inexact GM problems investigated in this paper are relatively small in size (e.g., graphs with less than 30 nodes). Hence, we choose a simple but effective classical solver–Integer Projected Fixed Point (IPFP) method [32] to solve Eq. (2). Briefly, IPFP solves a sequence of linear assignment problems using Taylor expansion around the current solution, then optimizes the original quadratic function along the ascent direction from the linear step, projecting back to feasible space if needed. In practice, IPFP converges after 5-10 iterations.

Let $\mathcal{G}_2$ be the imitating embodiment (learner) having the configuration $\boldsymbol{q}$ with associated forward kinematics map to compute $\mathcal{V}_2(\boldsymbol{q})$. We write $\mathcal{G}_2(\boldsymbol{q})$ to reflect that the node set depends on $\boldsymbol{q}$. The target embodiment $\mathcal{G}_1$ (expert) can be constructed from a controlled robot or uncontrolled human embodiment estimated by some human skeleton models [33]. We assume that $n_2 \leq n_1$ such that every node in $\mathcal{G}_2$ corresponds to one node in $\mathcal{G}_1$. Given the optimal correspondence $\boldsymbol{X}^*$ solution of the inexact GM problem Eq. (1), the embodiment divergence is defined as the sum of corresponding global feature distances, i.e., the distance sum of corresponded joint poses between embodiments

$$d_C(\mathcal{G}_1, \mathcal{G}_2(\boldsymbol{q})) = \text{tr}(\boldsymbol{X}^* \boldsymbol{D}^\intercal)$$
$$\text{s.t. } \boldsymbol{D}_{i,j} = d^v(\boldsymbol{T}_1^i, \boldsymbol{T}_2^j) \, \forall i \in \{1, \ldots, n_1\}, \forall j \in \{1, \ldots, n_2\}, \text{tr}(\cdot) \text{ is the trace operator.} \tag{4}$$

Note that $d_C(\mathcal{G}_1, \mathcal{G}_2)$ is a divergence on the space of all geometric embodiment graphs $\mathfrak{G}$, given the node sets as empirical distributions. To see this, consider any $\mathcal{G}_1, \mathcal{G}_2 \in \mathfrak{G}$, Eq. (4) can be written as a sum of distances between embodiment nodes $d_C(\mathcal{G}_1, \mathcal{G}_2) = \sum_{i,j} \boldsymbol{X}_{i,j}^* d^v(\boldsymbol{T}_1^i, \boldsymbol{T}_2^j)$. Thus, d_C satisfies non-negativity due to the metric d^v. Moreover, by construction, $d_C(\mathcal{G}_1, \mathcal{G}_2) = 0$ if and only if $\mathcal{G}_1 = \mathcal{G}_2$, hence satisfying positivity property.

3 Applications

We calibrate and fix the correspondence $\boldsymbol{X}^*$ by solving Eq. (1) at the initial embodiment configurations, e.g., rest joints of manipulators or default keyframe configuration of robot descriptions.

Riemannian Motion Policy. Provided that the optimal correspondence $\boldsymbol{X}^*$ is solved as described in previous sections, we propose a simple formulation describing reactive correspondence policy as a RMP between the leader Panda and the follower UR10 arms

$$\boldsymbol{f}(\boldsymbol{q}, \dot{\boldsymbol{q}}) = -\boldsymbol{K}_p \nabla_{\boldsymbol{q}} d_C - \boldsymbol{K}_v \dot{\boldsymbol{q}} \text{ with the weight matrix } \boldsymbol{M}(\boldsymbol{q}) = (a\,\|\nabla_{\boldsymbol{q}} d_C\| + \epsilon)\boldsymbol{I} \tag{5}$$

where $a > 0$ and $\nabla_{\boldsymbol{q}} d_C(\mathcal{G}_1, \mathcal{G}_2(\boldsymbol{q}))$ is the correspondence gradient. $\boldsymbol{f}$ is a PD controller with $\boldsymbol{K}_p, \boldsymbol{K}_v \succ 0$ gains. The Riemannian matrix is designed to be less relevant as the target $\boldsymbol{q}^*$ is approached (i.e., the decreasing gradient magnitude). Other RMPs such as self-collision avoidance, obstacle avoidance, and joint-limit avoidance specific to the controlling robot embodiment are designed similarly to [26]. The control rate of UR10 combined policy achieves approximately 100Hz, implying teleoperation applications.

Motion Retargeting for Inverse RL. We use human motion capture datasets and descriptions from [34] and perform motion retargeting to Unitree H1 for the walking task. We match the expert trajectory frequency with Unitree H1 control frequency and the episodic reward is formulated as the exponential of correspondence divergence summed over episode horizon $\sum_{t=0}^{T} \exp(-d_C(\mathcal{G}_1^t, \mathcal{G}_2(\boldsymbol{q}_t))/\sigma)$ with reward parameter $\sigma > 0$. The Unitree H1 PPO policy achieves approximately 10 seconds of walking. Note that the reward is purely correspondence divergence for matching the gait, and not specifically designed for locomotion.

4 Discussion & Conclusion

We present an approach to addressing the correspondence problem in imitation through the lens of inexact graph matching. Representing robotic and human embodiments as geometric graphs enables addressing the correspondence problem by formulating it as Lawler's Quadratic Assignment Problem (QAP) between embodiments with dissimilar morphologies, eliminating the need for a learning-based approach. However, the reliance on predefined feature functions for node and edge comparisons may limit its applicability to domains where such features are difficult to define. Additionally, the scalability of the graph-matching algorithm may become a bottleneck as the size of the graph increases significantly, although this was not a limiting factor in our experiments. Overall, our work suggests some immediate future works of kinematics teleoperations or imitation learning, such as the demonstration reuse of various embodiment datasets [22].

References

[1] S. Schaal, "Is imitation learning the route to humanoid robots?" *Trends in cognitive sciences*, vol. 3, no. 6, pp. 233–242, 1999.

[2] P. Abbeel, A. Coates, and A. Y. Ng, "Autonomous helicopter aerobatics through apprenticeship learning," *The International Journal of Robotics Research*, vol. 29, no. 13, pp. 1608–1639, 2010.

[3] J. Kober and J. Peters, "Imitation and reinforcement learning," *IEEE RAM*, 2010.

[4] P. Kormushev, S. Calinon, and D. G. Caldwell, "Imitation learning of positional and force skills demonstrated via kinesthetic teaching and haptic input," *Advanced Robotics*, vol. 25, no. 5, pp. 581–603, 2011.

[5] A. Boularias, J. Kober, and J. Peters, "Relative entropy inverse reinforcement learning," in *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, 2011, pp. 182–189.

[6] S. Calinon, F. D'halluin, E. L. Sauser, D. G. Caldwell, and A. G. Billard, "Learning and reproduction of gestures by imitation," *IEEE Robotics & Automation Magazine*, vol. 17, no. 2, pp. 44–54, 2010.

[7] T. Asfour, P. Azad, F. Gyarfas, and R. Dillmann, "Imitation learning of dual-arm manipulation tasks in humanoid robots," *International Journal of Humanoid Robotics*, vol. 5, no. 02, pp. 183–202, 2008.

[8] M. Lopes, F. S. Melo, and L. Montesano, "Affordance-based imitation learning in robots," in *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2007, pp. 1015–1021.

[9]   N. Ratliff, J. A. Bagnell, and S. S. Srinivasa, "Imitation learning for locomotion and manipulation," in *2007 7th IEEE-RAS International Conference on Humanoid Robots*.   IEEE, 2007, pp. 392–397.

[10] R. Chalodhorn, D. B. Grimes, K. Grochow, and R. P. Rao, "Learning to walk by imitation in low-dimensional subspaces," *Advanced Robotics*, vol. 24, no. 1-2, pp. 207–232, 2010.

[11] T. Osa, J. Pajarinen, G. Neumann, J. A. Bagnell, P. Abbeel, J. Peters *et al.*, "An algorithmic perspective on imitation learning," *Foundations and Trends® in Robotics*, vol. 7, no. 1-2, pp. 1–179, 2018.

[12] A. Billard, S. Calinon, and R. Dillmann, "Learning from humans," *Springer Handbook of Robotics*, pp. 1995–2014, 01 2016.

[13] T. Yu, C. Finn, A. Xie, S. Dasari, T. Zhang, P. Abbeel, and S. Levine, "One-shot imitation from observing humans via domain-adaptive meta-learning," *arXiv preprint arXiv:1802.01557*, 2018.

[14] K. Kim, Y. Gu, J. Song, S. Zhao, and S. Ermon, "Domain adaptive imitation learning," in *International Conference on Machine Learning*.   PMLR, 2020, pp. 5286–5295.

[15] A. Fickinger, S. Cohen, S. Russell, and B. Amos, "Cross-domain imitation learning via optimal transport," *arXiv preprint arXiv:2110.03684*, 2021.

[16] Y. Ma, A. Shen, D. Jayaraman, and O. Bastani, "Versatile offline imitation from observations and examples via regularized state-occupancy matching," in *ICML*, 2022.

[17] A. Gupta, C. Devin, Y. Liu, P. Abbeel, and S. Levine, "Learning invariant feature spaces to transfer skills with reinforcement learning," *arXiv preprint arXiv:1703.02949*, 2017.

[18] Y. Liu, A. Gupta, P. Abbeel, and S. Levine, "Imitation from observation: Learning to imitate behaviors from raw video via context translation," in *2018 IEEE international conference on robotics and automation (ICRA)*.   IEEE, 2018, pp. 1118–1125.

[19] T. Gangwani and J. Peng, "State-only imitation with transition dynamics mismatch," *arXiv preprint arXiv:2002.11879*, 2020.

[20] D. S. Raychaudhuri, S. Paul, J. Vanbaar, and A. K. Roy-Chowdhury, "Cross-domain imitation from observations," in *ICML*, 2021.

[21] T. Franzmeyer, P. Torr, and J. F. Henriques, "Learn what matters: cross-domain imitation learning with task-relevant embeddings," *Advances in Neural Information Processing Systems*, vol. 35, pp. 26 283–26 294, 2022.

[22] A. O'Neill, A. Rehman, A. Gupta, A. Maddukuri, A. Gupta, A. Padalkar, A. Lee, A. Pooley, A. Gupta, A. Mandlekar *et al.*, "Open x-embodiment: Robotic learning datasets and rt-x models," *arXiv preprint arXiv:2310.08864*, 2023.

[23] A. Khazatsky, K. Pertsch, S. Nair, A. Balakrishna, S. Dasari, S. Karamcheti, S. Nasiriany, M. K. Srirama, L. Y. Chen, K. Ellis *et al.*, "Droid: A large-scale in-the-wild robot manipulation dataset," *arXiv preprint arXiv:2403.12945*, 2024.

[24] H.-S. Fang, H. Fang, Z. Tang, J. Liu, C. Wang, J. Wang, H. Zhu, and C. Lu, "Rh20t: A comprehensive robotic dataset for learning diverse skills in one-shot," in *IEEE ICRA*, 2024.

[25] M. J. Kim, K. Pertsch, S. Karamcheti, T. Xiao, A. Balakrishna, S. Nair, R. Rafailov, E. Foster, G. Lam, P. Sanketi *et al.*, "Openvla: An open-source vision-language-action model," *arXiv preprint arXiv:2406.09246*, 2024.

[26] C.-A. Cheng, M. Mukadam, J. Issac, S. Birchfield, D. Fox, B. Boots, and N. Ratliff, "Rmpflow: A computational graph for automatic motion policy generation," in *WAFR*.   Springer, 2018.

[27] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *arXiv preprint arXiv:1707.06347*, 2017.

[28] J. Yan, X.-C. Yin, W. Lin, C. Deng, H. Zha, and X. Yang, "A short survey of recent advances in graph matching," in *Proceedings of the 2016 ACM on International Conference on Multimedia Retrieval*, 2016, pp. 167–174.

[29] J. Sola, J. Deray, and D. Atchuthan, "A micro lie theory for state estimation in robotics," *arXiv preprint arXiv:1812.01537*, 2018.

[30] F. Zhou and F. De la Torre, "Factorized graph matching," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*.   IEEE, 2012, pp. 127–134.

[31] J. Yan, S. Yang, and E. R. Hancock, "Learning for graph matching and related combinatorial optimization problems," in *IJCAI*, 2020.

[32] M. Leordeanu, M. Hebert, and R. Sukthankar, "An integer projected fixed point method for graph matching and map inference," *Advances in neural information processing systems*, vol. 22, 2009.

[33] C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja, M. Hays, F. Zhang, C.-L. Chang, M. G. Yong, J. Lee *et al.*, "Mediapipe: A framework for building perception pipelines," *arXiv preprint arXiv:1906.08172*, 2019.

[34] F. Al-Hafez, G. Zhao, J. Peters, and D. Tateo, "Locomujoco: A comprehensive imitation learning benchmark for locomotion," in *6th Robot Learning Workshop, NeurIPS*, 2023.