

A Nonparametric Motion Flow Model for Human Robot Cooperation

Sungjoon Choi, Kyungjae Lee, H. Andy Park, and Songhwai Oh

Abstract—In this paper, we present a novel nonparametric motion flow model that effectively describes a motion trajectory of a human and its application to human robot cooperation. To this end, motion flow similarity measure which considers both spatial and temporal properties of a trajectory is proposed by utilizing the mean and variance functions of a Gaussian process. We also present a human robot cooperation method using the proposed motion flow model. Given a set of interacting trajectories of two workers, the underlying reward function of cooperating behaviors is optimized by using the learned motion description as an input to the reward function where a stochastic trajectory optimization method is used to control a robot. The presented human robot cooperation method is compared with the state-of-the-art algorithm, which utilizes a mixture of interaction primitives (MIP), in terms of the RMS error between generated and target trajectories. While the proposed method shows comparable performance with the MIP when the full observation of human demonstrations is given, it shows superior performance when partial trajectory information is given.

I. INTRODUCTION

Traditionally, robots has been deployed to perform relatively simple and repetitive tasks in structured environments, where robots rarely interacted with humans. Recently, however, the growing need to reduce human workloads and risks and the advances in sensing, actuation, and computing capabilities have led to the efforts towards human-robot cooperation [1]. In order for a robot to be effectively deployed at a working space, where a human and a robot coexist in a close proximity, two main issues should be properly handled. The first issue is about inferring and recognizing the motion of a human co-worker and the second one is about teaching a robot how to act appropriately based on the inferred human motions.

In this paper, we mainly focus on the first problem of recognizing human motions by finding a mapping from a motion trajectory to a latent vector that effectively describes the trajectory. To this end, a nonparametric motion flow similarity measure is proposed by measuring the closeness of two trajectories in terms of both spatial distances and temporal flow directions. The proposed similarity measure alleviates existing time-alignment restrictions for computing

both spatial and temporal distances using the mean and variance functions of a Gaussian process [28]. This alignment-free property plays an important role in human robot cooperation tasks when it comes to recognizing the human motions given partial trajectories, i.e., early recognition. Furthermore, we present learning-based human robot cooperation based on the proposed motion flow model combined with inverse reinforcement learning [10] and stochastic trajectory optimization [11].

Both issues of recognizing the human motion and planning collaborative robot behavior have been widely investigated over the years. With respect to the first issue of the human motion inference, a number of studies have been made [3]–[5]. In [4], an interacting multiple model filtering approach is applied to infer the intended goal position of a human arm. [5] proposed a framework for modeling human motions by first clustering human demonstrations and finding the underlying reward function of each cluster using inverse reinforcement learning. While the approach of clustering human demonstrations is similar to ours, it mainly focused on identifying human user types.

Once human motion is properly recognized, a timely action of a robot should be generated. In [6], the inferred human motion is used to predict future movements where robot’s motion is planned by minimizing the penetration cost with a human. A pioneering work on generating interacting behaviors using interaction primitives (IPs), which is a special type of dynamic motor primitives [7], was proposed in [8]. IPs first parameterize trajectories of both the human hand and the robot state and the joint distribution of both parameters is modeled by a Gaussian distribution. Then, the interaction between a human and a robot is modeled by a conditional Gaussian model in the trajectory parameter space. The interaction primitives are extended to a mixture of interaction primitives (MIP) [9] using a Gaussian mixture model and successfully modeled multiple collaboration tasks between a human and robot.

The main contribution of this paper is to present a nonparametric motion flow model that can effectively map a trajectory to a latent vector considering both spatial and temporal aspects of a trajectory. Time-alignment as well as temporal adjustment, e.g., dynamic time warping, are not required for the proposed model. Furthermore, we present a human robot cooperation algorithm based on the proposed motion flow model and demonstrated its performance in both simulated and real-world environments.

II. RELATED WORK

Recently, a number of human motion descriptors have been proposed in both fields of robotics and computer vision. They can be grouped into four different categories based

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (NRF-2017R1A2B2006136), by the Next-Generation Information Computing Development Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (2017M3C4A7065926), and by the Brain Korea 21 Plus Project in 2017.

S. Choi is with the Kakao Brain, Korea (email: sam.choi@kakaobrain.com). K. Lee and S. Oh are with the Department of Electrical and Computer Engineering and ASRI, Seoul National University, Seoul 08826, Korea (e-mail: {kyungjae.lee and songhwai.oh}@cpslab.snu.ac.kr). H. Park is with Rethink Robotics, Boston, MA, USA (e-mail: apark@rethinkrobotics.com). This research was conducted during Sungjoon’s Ph.D. study in Seoul National University, Korea (sungjoon.choi@cpslab.snu.ac.kr).

TABLE I: Classification of motion description algorithms.

	Time-aligned	Alignment-free
Image-based	[18]–[21], [23]	
Trajectory-based	[12]–[15]	[15], [17], ours

on two different criteria. The first criterion is the source of information: a sequence of images or a sequence of positions. The second criterion is the ability to cope with partially observed trajectories: time-aligned and alignment-free. While some tasks will require explicit time-alignment of trajectories, e.g., tasks with timing restrictions or sequential operations, with respect to human robot collaborations, alignment-free descriptors have advantages in that it enables more natural and rapid interactions between a human and robot.

In [12]–[15], a sequence of positions of a human hand or an end effector of a robot is given as motion descriptors, which is often referred to as motion trajectories. Geometric invariant descriptions such as curvature and torsion of a motion trajectory are proposed in [12]–[14]. In [12], curvature, torsion and their first order derivatives with respect to arc-length parameters represent a motion description. In [14], integral invariant features are proposed. [15] proposed self similarity motion descriptions by constructing a self similarity matrix based on sigmoid distances between all pairs of points in a motion trajectory. [12]–[15] require additional dynamic time warping [16] to handle the time alignment problem. On the contrary, in [15], the feature for a motion trajectory are coefficients of the Fourier transform of the trajectory, alleviating the time alignment problem. [17] utilized the relative joint angles of each skeleton of a motion trajectory using the generalized Hough transform. The time invariant property is achieved as each instance of a motion trajectory votes for the motion class of the full trajectory.

In computer vision, spatial and temporal motion descriptions based on image information have been widely investigated [18]–[22]. [18] proposed four channel motion descriptors based on optical flow information and an associated similarity measure computed from the normalized correlation between two optical flows. In [19], a representative set of action prototypes is learned by training a binary tree using a motion description proposed in [18]. [21] presented a motion descriptor using the brightness gradients of space time interest points. In [22], the motion description is extracted by fitting B-spline surface on detected spatiotemporal salient points in image sequences.

III. PRELIMINARIES

In this section, we briefly review density matching reward learning (DMRL) [10] and Gaussian random paths (GRPs) [11]. The DMRL is used to learn the reward function of a robot whose input is a motion descriptor of a human co-worker trained from the proposed motion flow model for human robot cooperation. Once the reward function is learned properly, GRPs are used to generate an appropriate trajectory of the arm of a collaborating robot.

A. Density Matching Reward Learning

Density matching reward learning (DMRL) is a model-free reward learning algorithm [10], which finds the underlying reward function given expert’s demonstrations. Suppose that the estimated density $\hat{\mu}(s, a)$ of state s and action a is given, DMRL finds a reward function $R(\cdot)$ as follows:

$$\begin{aligned} & \underset{R}{\text{maximize}} \quad V(R) = \langle \hat{\mu}, R \rangle \\ & \text{subject to} \quad \|R\|_2 \leq 1, \end{aligned} \quad (1)$$

where the norm ball constraint $\|R\|_2 \leq 1$ is introduced to handle the scale ambiguity in the reward function [24] and $\langle \hat{\mu}, R \rangle = \int_{\mathcal{S} \times \mathcal{A}} \hat{\mu}(s, a) R(s, a) ds da$. For notational simplicity, we will denote x as a state-action pair, i.e., $x = (s, a)$.

To model a nonlinear reward function, kernel DMRL is proposed by assuming that the reward function has the following form:

$$\tilde{R}(x) = \sum_{i=1}^{N_U} \alpha_i k(x, x_i^U), \quad (2)$$

where $\{x_i^U\}_{i=1}^{N_U}$ is a set of N_U inducing points, $k(\cdot, \cdot)$ is a kernel function, and $\alpha \in \mathbb{R}^{N_U}$ determines the shape of the reward function. Then, we can reformulate (1) as the following unconstrained optimization:

$$\underset{\tilde{R}}{\text{maximize}} \quad \tilde{V} = \sum_{\forall x \in U} \hat{\mu}(x) \tilde{R}(x) - \frac{\lambda}{2} \|\tilde{R}\|_{\mathcal{H}}^2, \quad (3)$$

where λ controls the smoothness of the reward function, $U = \{x_i^U\}_{i=1}^{N_U}$ is a set of inducing points, and $\|\tilde{R}\|_{\mathcal{H}}^2$ is the squared Hilbert norm, which is often used as a regularizer for kernel machines.

B. Gaussian Random Paths

A Gaussian random path [11] defines a distribution over smooth paths, where a path $\mathbf{p}$ is defined as a vector-valued function and the domain is a time interval $\mathcal{I} \subseteq \mathbb{R}$ and the range is a set of locations. Given a kernel function $k(t, t')$ and a set of M anchoring locations $D_a = (\mathbf{t}_a, \mathbf{x}_a) = \{(t_i, x_i)\}_{i=1}^M$, a Gaussian random path $\mathcal{P}$ with a sequence of T test time indices $\mathbf{t}_{test} = \{t_i\}_{i=1}^T$ is specified by a mean path $\mu_{\mathcal{P}}$ and a covariance matrix $K_{\mathcal{P}}$, i.e., $\mathcal{P} \sim \mathcal{N}(\mu_{\mathcal{P}}, K_{\mathcal{P}})$, where

$$\begin{aligned} \mu_{\mathcal{P}} &= \mathbf{k}(\mathbf{t}_a, \mathbf{t}_{test})^T (\mathbf{K}_a + \sigma_w^2 I)^{-1} \mathbf{x}_a, \\ K_{\mathcal{P}} &= \mathbf{K}_{test} - \mathbf{k}(\mathbf{t}_a, \mathbf{t}_{test})^T (\mathbf{K}_a + \sigma_w^2 I)^{-1} \mathbf{k}(\mathbf{t}_a, \mathbf{t}_{test}), \end{aligned} \quad (4)$$

$\mathbf{k}(\mathbf{t}_{test}, \mathbf{t}_a) \in \mathbb{R}^{T \times M}$ is a kernel matrix of test time indices and anchoring time indices, $\mathbf{K}_a = \mathbf{K}(\mathbf{t}_a, \mathbf{t}_a) \in \mathbb{R}^{M \times M}$ is a kernel matrix of anchoring time indices, and $\mathbf{K}_{test} = \mathbf{K}(\mathbf{t}_{test}, \mathbf{t}_{test}) \in \mathbb{R}^{T \times T}$ is a kernel matrix of test time indices. In this paper, a squared exponential kernel function is used as the kernel function. An ϵ run-up method was further proposed to add a directional constraint on the path distribution [11].

Gaussian random paths are used for two purposes. First, as the GRP defines a distribution over paths, diverse smooth paths interpolating a set of anchoring points can be efficiently sampled. While existing trajectory optimization methods,

such as STOMP [25] and CHOMP [26], can be also used, we have found that trajectory optimization with GRPs is faster and produces a higher quality solution on problems that are sensitive to the choice of an initial trajectory [11]. Secondly, the mean path of GRPs can be used in a preprocessing step of the proposed method for smoothing and computing time derivatives.

IV. NONPARAMETRIC MOTION FLOW MODEL

In this section, we propose a nonparametric motion flow model for inferring and recognizing human motion trajectories, where a motion flow is defined as a mapping between positions of a trajectory to its derivatives. Given motion trajectories of a human, we first cluster trajectories using spectral clustering with the proposed motion flow similarity measure, where trajectories of each cluster are used to train a nonparametric motion flow model. Finally, the motion description of a trajectory is achieved by computing how close the trajectory is to each motion flow model in terms of the motion flow similarity measure.

In the following sections, we will focus on the position of the right hand of a human, however, higher dimensional inputs, e.g., joint positions of a right arm, can be also used. Suppose that $\mathbf{x} \in \mathbb{R}^3$ and $\dot{\mathbf{x}} \in \mathbb{R}^3$ are the position and velocity of a right hand. Then, the flow function $f(\cdot)$ maps $\mathbf{x}$ to $\dot{\mathbf{x}}$, i.e., $f: \mathbf{x} \mapsto \dot{\mathbf{x}}$.

For notational simplicity, we will denote ξ as a trajectory containing both positions and velocities, i.e., $\xi = \{(\mathbf{x}_k, \dot{\mathbf{x}}_k)\}_{k=1}^L$, where L is the length of a trajectory and $\mathbf{x}_{i,t}$ and $\dot{\mathbf{x}}_{i,t}$ are the position and velocity at time t for the i th trajectory ξ_i , respectively.

A. Motion Flow Similarity Measure

Suppose that two trajectories ξ_i and ξ_j are given. We present a nonparametric motion flow similarity measure based on both spatial and temporal aspects of trajectories, where the spatial similarity indicates how spatially close ξ_i and ξ_j are and the temporal similarity represents how much velocities of ξ_i and ξ_j are aligned.

Given two trajectories ξ_i and ξ_j , a motion flow similarity measure $d(\xi_i, \xi_j)$ of ξ_i and ξ_j is defined as:

$$d(\xi_i; \xi_j) = \frac{1}{L} \sum_{t=1}^L (d_{\cos}(\dot{\mathbf{x}}_{i,t}, \hat{\mu}_j(\mathbf{x}_{i,t})) + \hat{\sigma}_j^2(\mathbf{x}_{i,t})), \quad (5)$$

where $\hat{\mu}_j(\cdot)$ and $\hat{\sigma}_j^2(\cdot)$ are the Gaussian process mean function and variance function trained with input-output pairs $(\mathbf{x}_{j,t}, \dot{\mathbf{x}}_{j,t})$ from $\xi_j = \{(\mathbf{x}_{j,t}, \dot{\mathbf{x}}_{j,t})\}_{t=1}^L$. $\hat{\mu}_j(\cdot)$ and $\hat{\sigma}_j^2(\cdot)$ are defined as follows:

$$\hat{\mu}_j(\mathbf{x}_{i,t}) = k(\mathbf{x}_{i,t})^T (k(\mathbf{X}, \mathbf{X}) + \sigma_w^2 I)^{-1} \mathbf{y} \quad (6)$$

and

$$\hat{\sigma}_j^2(\mathbf{x}_{i,t}) = k_* - k(\mathbf{x}_{i,t})^T (k(\mathbf{X}, \mathbf{X}) + \sigma_w^2 I)^{-1} k(\mathbf{x}_{i,t}) \quad (7)$$

where $\mathbf{X} = \{\mathbf{x}_{j,t}\}_{t=1}^L$, $\mathbf{y} = \{\dot{\mathbf{x}}_{j,t}\}_{t=1}^L$, $k_* = k(\mathbf{x}_{i,t}, \mathbf{x}_{i,t})$, $k(\mathbf{x}_{i,t}) \in \mathbb{R}^n$ is a covariance vector between the test point $\mathbf{x}_{i,t}$ and L input points $\mathbf{X}$, and $k(\mathbf{X}, \mathbf{X}) \in \mathbb{R}^{L \times L}$ is a

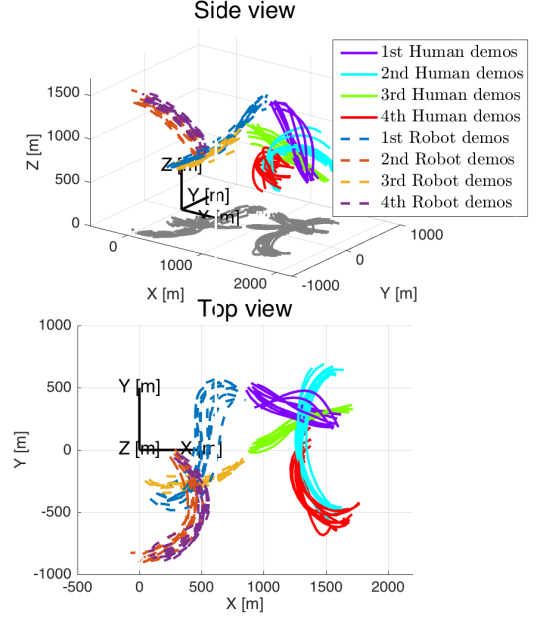


Fig. 1: A side view and top-down view of 40 trajectories clustered in different colors.

covariance matrix between L data points $\mathbf{X}$. Here, $d_{\cos}(\cdot, \cdot)$ is a cosine distance between two velocities:

$$d_{\cos}(\dot{\mathbf{x}}_a, \dot{\mathbf{x}}_b) = 1 - \frac{\dot{\mathbf{x}}_a^T \dot{\mathbf{x}}_b}{\|\dot{\mathbf{x}}_a\|_2 \|\dot{\mathbf{x}}_b\|_2}. \quad (8)$$

Intuitively speaking, (5) indicates a distance measure between two trajectories considering both spatial and temporal properties. To be more specific, the first term inside the summation in (5), i.e., $\frac{1}{L} \sum_{t=1}^L d_{\cos}(\dot{\mathbf{x}}_{i,t}, \hat{\mu}_j(\mathbf{x}_{i,t}))$, indicates the *temporal similarity* between ξ_i and ξ_j as it represents how the velocities of ξ_i and ξ_j are aligned. Similarly, the second term, i.e., $\frac{1}{L} \sum_{t=1}^L \hat{\sigma}_j^2(\mathbf{x}_{i,t})$, represents the *spatial similarity* between ξ_i and ξ_j since the Gaussian process variance increases as the distance between the test input and training samples increases, i.e., the predictive variance is high if there are no nearby training samples. Both similarities go to zeros if $\xi_i = \xi_j$. We would like to emphasize that time-alignments of ξ_i and ξ_j are not required as they are not compared in a point-wise manner. We would like to note that the proposed motion flow similarity measure can be applied to any input and output sequences by treating $\mathbf{x}$ as an input and $\dot{\mathbf{x}}$ as an output.

B. Training Motion Flow Model

Suppose that motion trajectories are given as training data. The collected trajectories are clustered using a spectral clustering algorithm [27]. The adjacency matrix $A \in \mathbb{R}^{N \times N}$ is computed from the similarity measure (5), where N is the number of trajectories. In particular, $[A]_{i,j} = d(\xi_i, \xi_j)$ and $A = \frac{1}{2} (\bar{A}^T + \bar{A})$ for ensuring positive definiteness of the adjacency matrix. Clustering results of four clusters using 40 training samples are shown in Figure 1 (see Section V-A for information about how samples are collected).

Once trajectories are clustered, a motion flow model is trained using Gaussian process regression (GPR) [28]. In

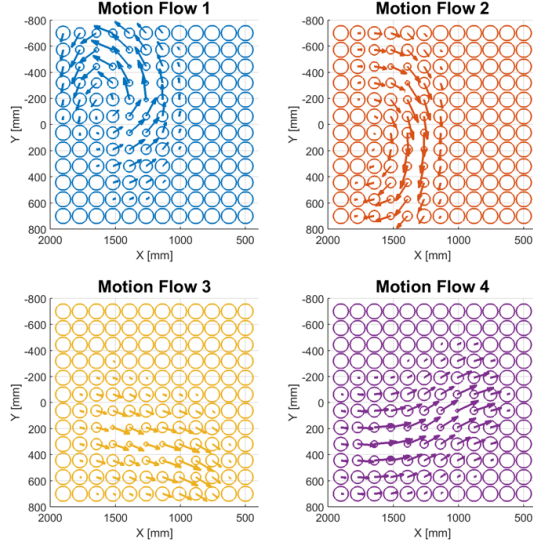


Fig. 2: A top-down view of nonparametric motion flow models. A motion flow is shown with an arrow and the size of a circle indicates the predicted variance $\hat{\sigma}^2(\cdot)$ at each grid.

particular, the motion flow function of each cluster is trained using GPR, where the positions of trajectories are given as input training data and corresponding velocities are given as output training data. In other words, given K clusters, we compute K GP mean functions, $\hat{\mu}_k(\cdot)$, and variance functions, $\hat{\sigma}_k^2(\cdot)$. Furthermore, we compute K motion flow similarity measures

$$d_k(\xi_i) = \frac{1}{L} \sum_{t=1}^L (d_{\cos}(\dot{\mathbf{x}}_{i,t}, \hat{\mu}_k(\mathbf{x}_{i,t})) + \hat{\sigma}_k^2(\mathbf{x}_{i,t})). \quad (9)$$

The motion flow model trained with each cluster is illustrated in Figure 2, where the arrow indicates the Gaussian process mean functions $\hat{\mu}(\cdot)$ and the size of the circle is proportional to the predictive variance $\hat{\sigma}^2(\cdot)$. We can easily see that the direction of the arrow coincides with the direction of trajectories in each cluster while the predictive variances of GP increase as we move away from the mean path.

C. Motion Flow Description Inference

Finally, the trained motion flow model is used to extract motion flow descriptions of a motion trajectory. The motion flow descriptions are computed as follows. Suppose that the number of clusters is K . Then the motion flow description $\mathbf{p}$ of a motion trajectory ξ is a K -dimensional vector on a simplex, where k -th element is proportional to $\exp(-d_k(\xi)^2)$, i.e., $\{\mathbf{p} \in [0, 1]^K \mid p_k \propto \exp(-d_k(\xi)^2) \sum_{k=1}^K p_k = 1\}$. We will denote the mapping between a trajectory ξ and motion descriptor $\mathbf{p}$ as $\phi(\cdot)$, i.e., $\phi: \xi \mapsto \mathbf{p}$.

The motion flow description $\mathbf{p}$ representing a trajectory is used to model the reward function of a robot given interacting demonstrations consisting of motion trajectories of a human and robot. As the motion flow description uses flow functions trained with a GP, time alignment of trajectories is not required. This alignment-free property plays a significant

role in early recognition of the human motion in upcoming experiments shown in Section V.

V. EXPERIMENTS

In this section, the performance of the proposed nonparametric motion flow model is extensively validated through human robot cooperation. The proposed method is compared with a mixture of interaction primitives (MIP) [9] in terms of root mean squared (RMS) errors between predicted and target trajectories in the test set under different scenarios.

A. Collecting Interacting Demonstrations

Before presenting the experimental results of the proposed method, let us introduce the interacting demonstration dataset used throughout the experiments. We assume that a human and a robot are collaboratively arranging the space in a close proximity under four different modes:

- 1) Center-hand-over: If a human hands over an object in the middle, a robot reach the object by stretching its end-effector to the middle.
- 2) Right-hand-over: If a human hands over an object in the right, a robot reach the object by stretching its end-effector to the right.
- 3) Right-swipe: If a human swipes his hand to the right, a robot operates its own arranging movement of raising its end-effector to the top right.
- 4) Left-swipe: If a human swipes his hand to the left, a robot operates its own arranging movement of raising its end-effector to the top right.

We have collected interacting demonstrations between two persons. In particular, we have recorded three dimensional positions of right hands of two persons using a Vicon motion capture system at 10Hz. One person performs the role of a robot considering the working range of a robot.¹ We have collected 20 demonstrations for each mode and the collected demonstrations are further divided into two sets: train and test sets. The collected trajectories are shown in the left of Figure 3.

B. Implementing Human Robot Cooperation Algorithm

Once a nonparametric motion flow model is trained from demonstrations of a human, the reward function of the robot that explains the interacting behavior is optimized using inverse reinforcement learning. This indirect way of modeling the interacting behaviors of the robot is particularly important as it allows incorporating additional considerations, such as collision avoidance, to the optimized reward function. The reward function r is a function of both motion flow descriptions of a human $\phi(\xi^O)$ and the end-effector position of a robot $\mathbf{x}^R$, i.e., $r(\phi(\xi^O), \mathbf{x}^R)$.

Suppose that a set of N interacting demonstrations of both a human and a robot, $\mathcal{D} = \{(\xi_i^H, \xi_i^R)\}_{i=1}^N$, is given and the number of clusters in the motion flow model is K . For notational simplicity, we assume that the length of each trajectory is L . Let $\xi_{i,1:t}^H$ be the i th trajectory of a human from time 1 to t , where $1 \leq t \leq L$. We first extract motion flow features Φ from human demonstrations at each time

¹A Baxter robot from Rethink Robotics is considered in this paper.

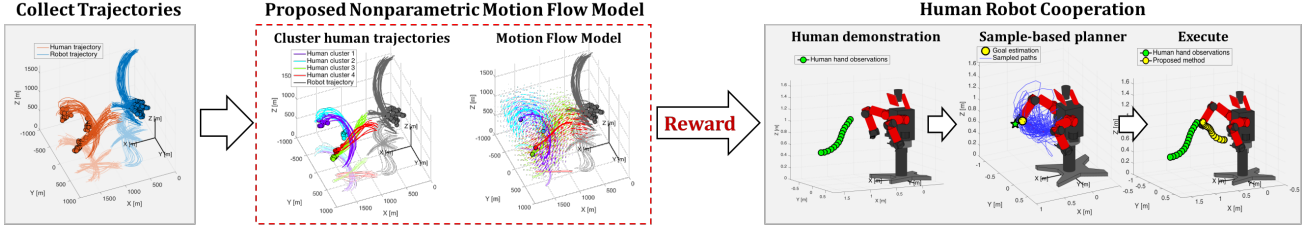


Fig. 3: An overview of the human robot cooperation algorithm based on the proposed motion flow model. Given interacting demonstrations between a human and robot, a motion flow model is trained by first clustering motion trajectories of a human and compute motion flow similarity measures of each cluster. Then, the underlying reward function of a robot is optimized and further used in the execution phase using a sampling-based trajectory optimizer.

step, i.e., $\Phi = \{\phi(\xi_{t,1:t}^H) | t = 1, \dots, L\}_{i=1}^N$ along with the position of the end-effector of a robot. In other words, we extract $N \times L$ pairs of K dimensional motion flow features and corresponding end-effector positions from $\mathcal{D}$, which will be used as a train set. Density matching reward learning [10] is used to optimize the underlying reward function of a robot given recognized human motions. However, other inverse reinforcement learning algorithms, e.g., Gaussian process inverse reinforcement learning [29] or continuous inverse optimal control [30], can also be used.

Furthermore, we similarly define the reward of a trajectory of the end-effector of a robot by averaging the reward of each position in the trajectory. Let ξ^R be the end-effector trajectory of a robot, then the reward of the robot trajectory given the observed human trajectory ξ^O is defined as:

$$R(\xi^R | \xi^O) = \frac{1}{L} \sum_{t=1}^L r(\phi(\xi_t^O), \mathbf{x}_t^R). \quad (10)$$

Once the motion of a human is inferred, a pertinent trajectory of the right arm of a robot is optimized with respect to the reward function. First, we estimate the final pose and its time derivative of an end-effector trajectory. Suppose that $\mathcal{D} = \{(\xi_i^H, \xi_i^R)\}_{i=1}^N$ is the interacting demonstrations of a human and a robot. Then, the final pose of a robot $\hat{\mathbf{x}}_L^R$ given an observed human trajectory ξ^O is estimated as:

$$\hat{\mathbf{x}}_L^R = \sum_{i=1}^N \mathbf{x}_{i,L}^R \frac{p(\xi_i^O | \xi_i^H)}{\sum_{j=1}^N p(\xi_j^O | \xi_j^H)}, \quad (11)$$

where L is the length of a trajectory and $\mathbf{x}_{i,L}^R$ is the last position of the i th robot trajectory ξ_i^R . The final time derivative $\hat{\dot{\mathbf{x}}}_L^R$ is similarly estimated by replacing $\mathbf{x}_{i,L}^R$ to $\dot{\mathbf{x}}_{i,L}^R$ in (11).

Finally, the interacting trajectory of a robot end-effector, $\hat{\xi}^R$, is computed as:

$$\hat{\xi}^R = \sum_{i=1}^K \xi_i^R \frac{\exp(R(\xi_i^R | \xi^O))}{\sum_{j=1}^K \exp(R(\xi_j^R | \xi^O))}, \quad (12)$$

where $R(\xi_i^R | \xi^O)$ is the reward of the trajectory of a robot's end-effector in (10) and ξ_i^R is the i th sampled trajectory drawn from a Gaussian random path distribution $p_{GRP}(\xi^R)$. The anchoring points of the GRP consist of $\{(0, \mathbf{x}_t^R), (1 - \epsilon, \hat{\mathbf{x}}_L^R - \epsilon \dot{\mathbf{x}}_L^R), (1, \hat{\mathbf{x}}_L^R)\}$, where $\mathbf{x}_t^R$ is the current end-effector position and the second anchoring point is added

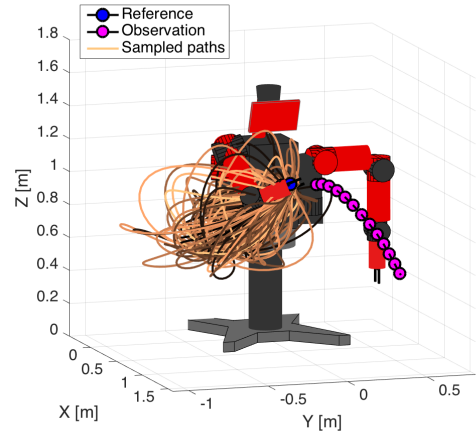


Fig. 4: Sampled paths for a Baxter robot using Gaussian random paths.

to incorporate the estimated final heading of a robot, where ϵ is set to 0.01. 200 paths of the end-effector of a Baxter robot sampled from the GRP are shown in Figure 4. The path sampling is done in the 7-DoF configuration space of a Baxter's right arm and the three dimensional paths of the end-effector are computed using the forward kinematics. The overall computation took less than 100ms in MATLAB on a 2.2GHz quad-core processor. We will refer the human robot cooperation algorithm with the proposed motion flow model as intention aware apprenticeship learning (IAAL). The overview of IAAL is shown in Figure 3.

C. Planning With Partial Observations

In this experiment, we validated the early recognition performance of IAAL compared to MIL by varying the observation ratio of a human hand trajectory. We vary the observation ratio from 0.2 to 1.0, where 0.2 indicates giving first 20% of the human hand trajectory. The number of clusters for the IAAL and MIP are set to five.

The prediction results are shown in Figure 5. While IAAL shows comparable performance with MIP given the full observation of a human hand trajectory, the prediction performance of MIP degenerates significantly as the observation ratio decreases. On the contrary, the performance drop of

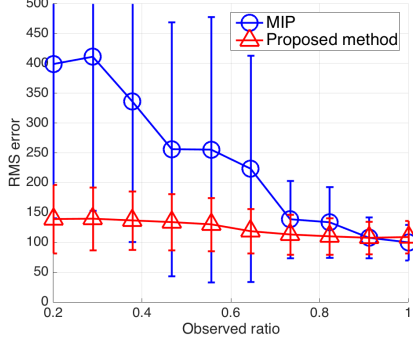


Fig. 5: RMS prediction errors of the proposed IAAL and MIP as a function of the observation ratio.

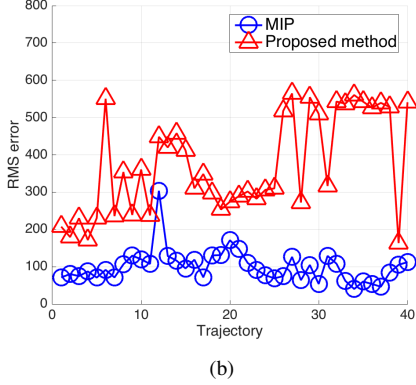
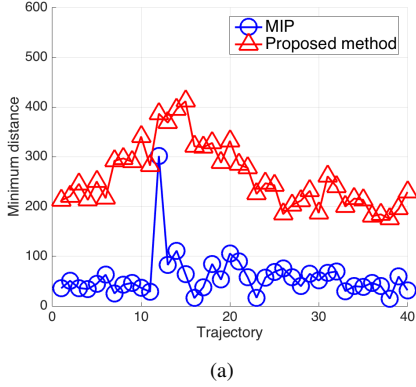


Fig. 6: (a) Minimum distances to the obstacle of each run. (b) RMS prediction errors of each run.

IAAL is minuscule reflecting the early prediction ability of IAAL. This is largely due to the proposed motion flow similarity measure, which is free from the time alignment of trajectories.

D. Planning With Obstacles

The objective of this experiment is to validate the ability of IAAL of incorporating additional considerations. In particular, an obstacle is assumed to be placed in the middle of the target trajectory of the end-effector. We consider not only the collision with the end-effector of a robot, but also three joints (wrist, elbow, and shoulder) in the right arm of a

Baxter robot. This becomes possible as we sample paths in the configuration (joint) space of the right arm of a Baxter robot and use forward kinematics to compute the trajectories of four joints (hand, wrist, elbow, and shoulder) in a three dimensional Cartesian coordinate.

The following log-barrier function is designed for collision avoidance:

$$C^{obs}(\xi^R) = -\log(\gamma \cdot (d_{min} - \alpha) + \epsilon) + \beta, \quad (13)$$

where d_{min} is the minimum distance from the obstacle to four trajectories of the right arm, i.e., right hand, wrist, elbow, and shoulder, $\alpha = 100$, $\beta = -12.8$, $\gamma = 5.73 \times 10^{-9}$, and $\epsilon = 10^{-6}$.

As MIP cannot directly handle additional constraints, only IAAL is modified by subtracting (13) to the reward function in the trajectory optimization process in (12) as follows:

$$\hat{\xi}^R = \sum_{i=1}^K \xi_i^R \frac{\exp(R(\xi_i^R|\xi^O) - C^{obs}(\xi_i^R))}{\sum_{j=1}^K \exp(R(\xi_j^R|\xi^O) - C^{obs}(\xi_j^R))}.$$

The minimum distance to obstacles and the RMS prediction error of 40 test runs, are shown in Figure 6(a) and 6(b), respectively. While the RMS prediction errors increase as the end-effector trajectories of IAAL detour to prevent collision, the minimum distance from obstacles to four trajectories of the right arm (end-effector, wrist, elbow, and shoulder) exceeds 180mm for all 40 test runs. Snapshots of a Baxter robot executing hand over tasks with and without obstacles are shown in Figure 7.

E. Manipulator-X Experiments

In this section, Manipulator-X from Robotis, a 7-DoF manipulator, is used to validate the applicability of the proposed algorithm in real-world environments. The interacting demonstrations in Section V-A is also used in this regard, where the positions of demonstrations are rescaled to compensate the smaller working range of Manipulator-X.

Once demonstrations are downscaled to fit Manipulator-X, the proposed IAAL algorithm is used plan a pertinent action of the manipulator, where the number of clusters is set to five. The human hand position is captured by a Vicon motion capture system and the manipulator re-plans its trajectory every two seconds with respect to the recognized human motion. The overall planning procedure took less than 400ms in MATLAB on a 2.2GHz quad-core processor and the manipulator halts its motion while planning.

We first conducted simple hand-over experiments, where an additional grasping motion is manually programmed. The final grasped positions along with the average trajectory in each cluster of the motion flow model are shown in Figure 8. Interestingly, the robot successfully reached and grasped the object in the regions, where the interacting demonstrations have not covered (see the green cylinders). However, if the location of the object is too far from the collected demonstrations, it fails (see the red cylinder).

The objective of the second set of experiments is to see how the robot reacts to varying intentions of a human. As the trajectory of Manipulator-X is re-planned every two seconds, it should be able to correctly modify the trajectory based on

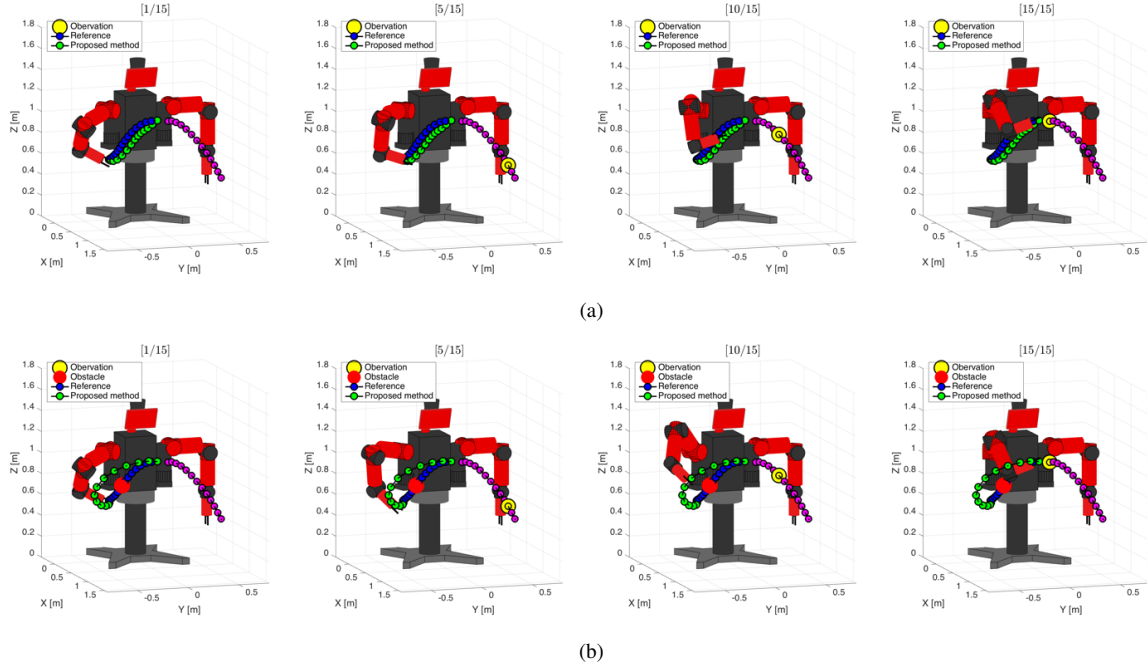


Fig. 7: (a) Snapshots of a Baxter robot executing the optimized trajectory using IAAL in a free space. (b) Snapshots of a Baxter robot executing the optimized trajectory using IAAL with an obstacle shown as a red circle in the middle of the target trajectory.

VI. CONCLUSION

In this paper, a motion flow model that describes motion trajectories of a human is proposed for human robot cooperation tasks. The main contribution of this paper is the nonparametric motion flow similarity measure based on both spatial and temporal similarities using a Gaussian process. The time-alignment of trajectories is not required for the proposed similarity measure and it plays an important role in early recognition of human motions. We presented a human robot cooperation algorithm based on the proposed motion flow model and compared it with a mixture of interaction primitives algorithm. The proposed algorithm has shown a superior performance with respect to the prediction error when partial trajectories of a human coworker are given.

REFERENCES

- [1] A. Bauer, D. Wollherr, and M. Buss, "Human-robot collaboration: a survey," *International Journal of Humanoid Robotics*, vol. 5, no. 01, pp. 47–66, 2008.
- [2] B. D. Argall, S. Chernova, M. Veloso, and B. Browning, "A survey of robot learning from demonstration," *Robotics and autonomous systems*, vol. 57, no. 5, pp. 469–483, 2009.
- [3] Y. Li and S. S. Ge, "Human-robot collaboration based on motion intention estimation," *IEEE/ASME Transactions on Mechatronics*, vol. 19, no. 3, pp. 1007–1014, 2014.
- [4] H. chaandar Ravichandar and A. Dani, "Human intention inference through interacting multiple model filtering," in *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*. IEEE, 2015, pp. 220–225.
- [5] S. Nikolaidis, R. Ramakrishnan, K. Gu, and J. Shah, "Efficient model learning from joint-action demonstrations for human-robot collaborative tasks," in *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 2015, pp. 189–196.

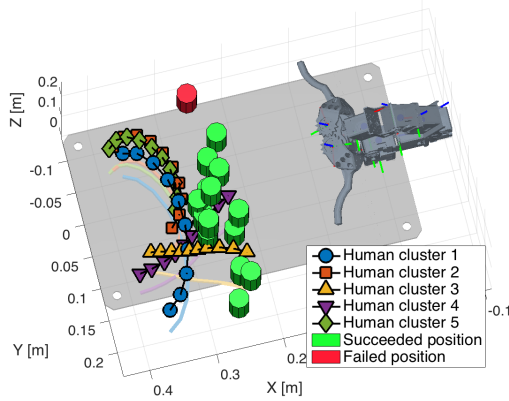


Fig. 8: Successful and failed final grasped positions of an object are shown with green and red cylinders, respectively. The mean trajectory of each cluster is shown with a different marker.

recognized motions. Figure 9 shows two cases: (1) the human co-worker changes his intention from left-swipe to center-hand-over; and (2) the intention is changed from center-hand-over to right-hand-over and back to center-hand-over. It shows that the proposed IAAL algorithm can successfully recognize the change in the user's motions and respond accordingly.

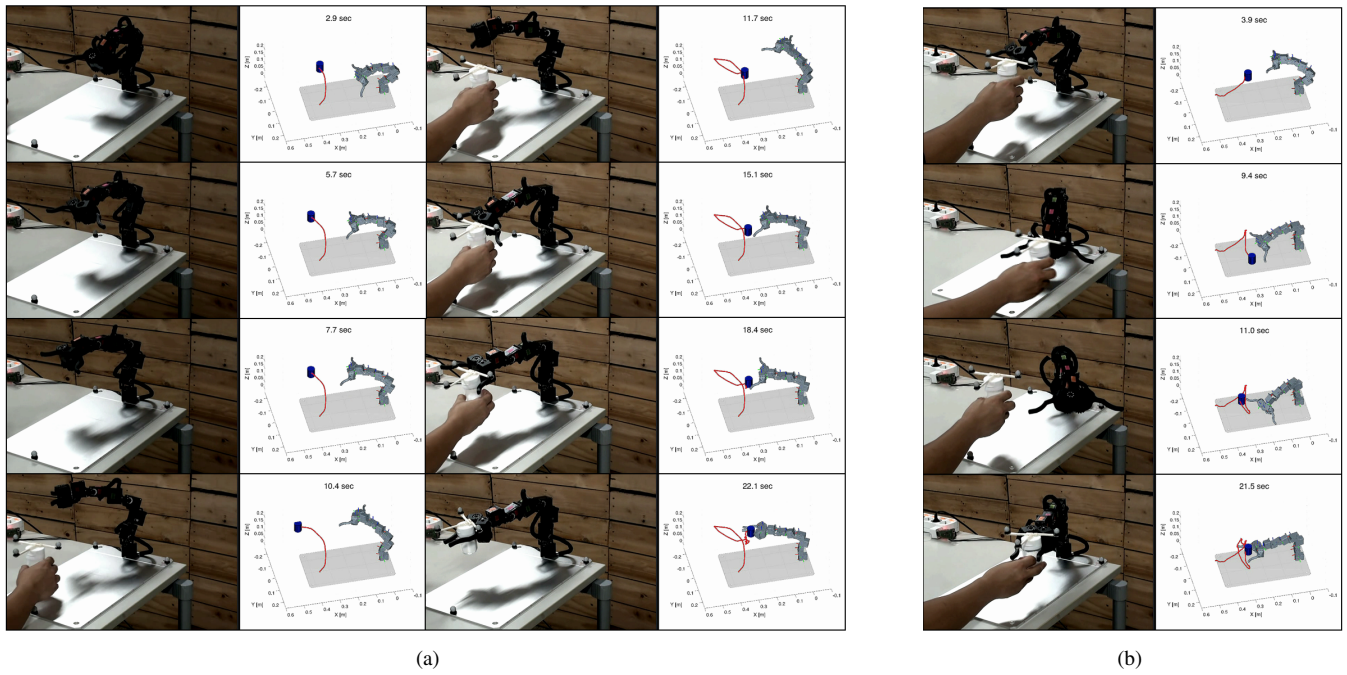


Fig. 9: Snapshots of experiments using Manipulator-X. (a) The intention is changed from left-swipe to center-hand-over. (b) The intention is changed from center-hand-over to right-hand-over and back to center-hand-over.

- [6] J. Mainprice and D. Berenson, "Human-robot collaborative manipulation planning using early prediction of human motion," in *IEEE Proc. of International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2013, pp. 299–306.
- [7] A. Paraschos, C. Daniel, J. R. Peters, and G. Neumann, "Probabilistic movement primitives," in *Advances in neural information processing systems (NIPS)*, 2013, pp. 2616–2624.
- [8] H. B. Amor, G. Neumann, S. Kamthe, O. Kroemer, and J. Peters, "Interaction primitives for human-robot cooperation tasks," in *IEEE Proc. of International Conference on Robotics and Automation (ICRA)*. IEEE, 2014, pp. 2831–2837.
- [9] M. Ewerton, G. Neumann, R. Lioutikov, H. B. Amor, J. Peters, and G. Maeda, "Learning multiple collaborative tasks with a mixture of interaction primitives," in *IEEE Proc. of International Conference on Robotics and Automation (ICRA)*. IEEE, 2015, pp. 1535–1542.
- [10] S. Choi, K. Lee, A. Park, and S. Oh, "Density matching reward learning," *arXiv preprint arXiv:1608.03694*, 2016.
- [11] S. Choi, K. Lee, and S. Oh, "Gaussian random paths for real-time motion planning," in *Proc. of the IEEE International Conference on Intelligent Robots and Systems (IROS)*, 2016.
- [12] S. Wu and Y. Li, "On signature invariants for effective motion trajectory recognition," *I. J. Robotics Res.*, vol. 27, no. 8, pp. 895–917, 2008.
- [13] J. Yang, Y. Li, K. Wang, Y. Wu, G. Altieri, and M. Scalia, "Mixed signature: an invariant descriptor for 3d motion trajectory perception and recognition," vol. 2012. Hindawi Publishing Corporation, 2011.
- [14] Z. Shao and Y. Li, "Integral invariants for space motion trajectory matching and recognition," *Pattern Recognition*, vol. 48, no. 8, pp. 2418–2432, 2015.
- [15] P. R. G. Harding and T. J. Ellis, "Recognizing hand gesture using fourier descriptors," in *17th International Conference on Pattern Recognition, ICPR 2004, Cambridge, UK, August 23-26, 2004.*, 2004, pp. 286–289.
- [16] K. K. Paliwal, A. Agarwal, and S. S. Sinha, "A modification over sakoe and chiba's dynamic time warping algorithm for isolated word recognition," in *IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP '82, Paris, France, May 3-5, 1982*, 1982, pp. 1259–1261.
- [17] D. Muench, W. Huebner, and M. Arens, "Generalized hough transform based time invariant action recognition with 3d pose information," in *SPIE Security+ Defence*. International Society for Optics and Photonics, 2014, pp. 92 530K–92 530K.
- [18] A. A. Efros, A. C. Berg, G. Mori, and J. Malik, "Recognizing action at a distance," in *9th IEEE International Conference on Computer Vision (ICCV 2003), 14-17 October 2003, Nice, France, 2003*, pp. 726–733.
- [19] Z. Lin, Z. Jiang, and L. S. Davis, "Recognizing actions by shape-motion prototype trees," in *IEEE 12th International Conference on Computer Vision, ICCV 2009, Kyoto, Japan, September 27 - October 4, 2009*, 2009, pp. 444–451.
- [20] A. Fathi and G. Mori, "Action recognition by learning mid-level motion features," in *2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2008), 24-26 June 2008, Anchorage, Alaska, USA, 2008*.
- [21] J. C. Niebles, H. Wang, and F. Li, "Unsupervised learning of human action categories using spatial-temporal words," *International Journal of Computer Vision*, vol. 79, no. 3, pp. 299–318, 2008.
- [22] A. Oikonomopoulos, M. Pantic, and I. Patras, "Sparse b-spline polynomial descriptors for human activity recognition," *Image Vision Comput.*, vol. 27, no. 12, pp. 1814–1825, 2009.
- [23] S. Samanta and B. Chanda, "Space-time facet model for human activity classification," *IEEE Trans. Multimedia*, vol. 16, no. 6, pp. 1525–1535, 2014.
- [24] A. Y. Ng and S. Russell, "Algorithms for inverse reinforcement learning," in *Proc. of the 17th International Conference on Machine Learning*, June 2000.
- [25] M. Kalakrishnan, S. Chitta, E. Theodorou, P. Pastor, and S. Schaal, "STOMP: Stochastic trajectory optimization for motion planning," in *Proc. of the International Conference of Robotics and Automation (ICRA)*. IEEE, 2011.
- [26] N. Ratliff, M. Zucker, J. A. Bagnell, and S. Srinivasa, "Chomp: Gradient optimization techniques for efficient motion planning," in *Proc. of the International Conference on Robotics and Automation (ICRA)*. IEEE, 2009.
- [27] A. Y. Ng, M. I. Jordan, Y. Weiss *et al.*, "On spectral clustering: Analysis and an algorithm," in *Advances in Neural Information Processing Systems (NIPS)*, vol. 2. MIT, 1998, 2002, pp. 849–856.
- [28] C. Rasmussen and C. Williams, *Gaussian processes for machine learning*. MIT press Cambridge, MA, 2006, vol. 1.
- [29] S. Levine, Z. Popovic, and V. Koltun, "Nonlinear inverse reinforcement learning with gaussian processes," in *Advances in Neural Information Processing Systems 24*. Curran Associates, Inc., December 2011.
- [30] S. Levine and V. Koltun, "Continuous inverse optimal control with locally optimal examples," in *International Conference on Machine Learning (ICML)*, 2013.