2012 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 7-12, 2012. Vilamoura, Algarve, Portugal

Feedback Motion Planning and Learning from Demonstration in Physical Robotic Assistance: Differences and Synergies

Martin Lawitzky José Ramón Medina Dongheui Lee Sandra Hirche
Institute of Automatic Control Engineering
Technische Universität München
80290 Munich, Germany
Email: {ml,medina,dhlee,hirche}@tum.de

***Abstract*— Goal-directed physical assistance to the human is one of the most challenging problems in the area of human-robot interaction. Planning and learning from demonstration represent two conceptually different approaches to achieve goal-directed behavior. Here we examine the properties of a planning-based and a learning-based approach in the context of physical robotic assistance for the prototypical task of cooperative object maneuvering. In order to exploit the complementary strengths of planning and learning-based approaches we derive three novel synergy strategies. The algorithms are experimentally evaluated in a human user study in a planar virtual-reality scenario and in a proof-of-concept study with a human-sized mobile robot with two 7DoF arms. The results show that combinations of planning and learning algorithms are superior over the individual approaches.**

I. Introduction

Physical robotic assistance has a high potential for application in various areas ranging from intelligent transport in domestic and industrial settings to mobility assistance for elderly. In order to assist beyond the functionality of a purely reactive follower as for example in [1], the robotic partner must exhibit *goal-directed* assistive behavior. Several methods to equip a robot with necessary task knowledge for physical manipulation assistance are explored in the literature. Most of these approaches fall into the category of either a) planning-based methods, or b) motion reproduction based on imitation learning, each having different advantages and disadvantages. To the best of the authors' knowledge, there are no works on the *combination* of planning and learning in physical human-robot interaction (pHRI), which is the focus of this paper.

Even motion planning for physical robotic assistants itself finds only little reference in the literature. In [2] a heuristic method to improve the sampling strategy of a path planner through haptic input is presented. Feedback motion planning for haptic guidance using a cell-decomposition method is proposed by [3] in the context of computer-aided design. As a result, a channel similar to *Virtual Fixtures* [4] is generated by a planning algorithm to guide a user in virtual reality from start to goal [5]. None of these methods, however, allow the human partner to deviate freely from the pre-planned path in order to suit the human's motion preference. Instead, the focus of these methods is the attraction to the distance-optimal path. In order to address this issue and the need for instant decision making due to the physical coupling, here we propose a feedback-planning approach inspired by [6].

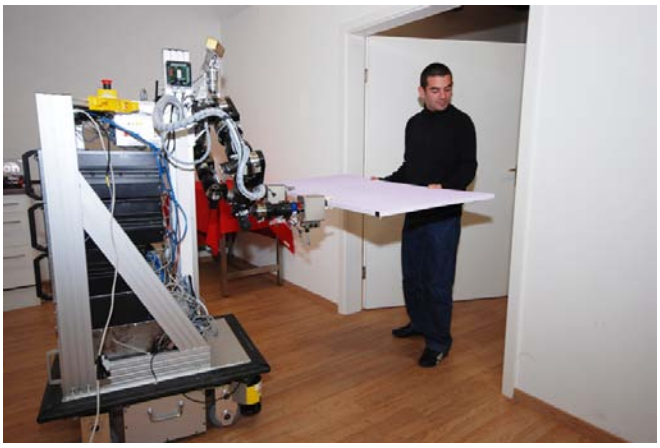


Fig. 1: Experimental setup with mobile manipulator used in evaluation

Within the programming-by-demonstration framework a significant body of methods to learn motion patterns from human demonstration, to recognize such, and even to estimate the human intention are developed e.g. in [7]–[10]. It is well-known that the estimation of the human partner's intention is crucial for intuitive physical human-robot interaction [11] and outperforms low-stiffness position tracking regarding subjective quality of assistance [12]. Programming-by-demonstration through a teleoperation system is explored by [7]. Interaction motion is reproduced based on previous observations on point-to-point level [8] and along a single dimension [9]. In our own earlier work [10] an incremental learning approach based on time-based hierarchical Hidden-Markov-Models is considered to estimate human intentions in more complex cooperative object maneuvring tasks. An exhaustive survey on general motion reproduction algorithms is provided in [13].

The contribution of this work is a novel approach to facilitate goal-directed behavior in physical human-robot interaction (pHRI) based on the synergetic combination of planning- and learning-based approaches. To this end we investigate the complementary properties of a learning-based motion prediction algorithm [10], [14] and a feedback-planning algorithm [6] in the context of pHRI, where the physical coupling between the robot and the human imposes particular challenges in terms of real-time capability, adaptability, and safety of the decision making scheme. Exploiting

the complementary strength of planning- and learning-based approaches we explore three different synergetic strategies based on i) plan-based initialization of the learning scheme, ii) homotopy-blending of planning and learning strategies based on the human motion prediction quality, and iii) a cost-based fusion of planning and learning strategies. As a prototypical task we investigate the cooperative transport of an object from an initial to a final configuration through a cluttered environment. We experimentally evaluate the novel approaches in a human user study in a planar virtual-reality setup and present an experimental proof-of-concept study in six degrees of freedom (DoF) with a human-sized mobile platform equipped with two 7-DoF anthropomorphic arms. The results show that combinations of planning and learning algorithms are superior over the individual approaches.

The remainder of this paper is organized as follows: Section II investigates planning and learning strategies in pHRI, which are then combined in novel synergetic strategies in Section III. The experimental setups described in Section IV are used to generate the results presented in Section V.

Notation: Bold characters are used for vectors and matrices. The configuration space of the manipulated rigid object is denoted $\mathcal{C}$ which is a manifold $\mathcal{C} = \mathbb{R}^2$ in two-dimensional case without rotation and $\mathcal{C} = \mathbb{R}^3 \times \mathbb{RP}^3$ in the three-dimensional case with rotation. The *obstacle region* is denoted $\mathcal{C}_{\text{obs}} \subseteq \mathcal{C}$. The leftover configurations are called *free space* which is denoted $\mathcal{C}_{\text{free}} = \mathcal{C} \setminus \mathcal{C}_{\text{obs}}$.

II. Planning and Learning in pHRI

Our studies of physical robotic assistants show that the quality of assistance benefits from a goal-oriented robot behavior in load transport tasks [15], [16]. In this context, plan-based [17] as well as motion reproduction strategies [10] successfully generate goal-oriented behaviors in full-scale cooperative transport tasks. In the following we will highlight the complementary strength of both approaches in order to derive suitable ways of combining them in a synergistic way. Therefore the general problem setting is presented first.

A. Problem setting

We consider the problem of a human-robot dyad cooperatively moving a rigid object from an initial configuration $\boldsymbol{x}(0) \in \mathcal{C}_{\text{free}}$ to a final configuration $\boldsymbol{x}(t_{\text{goal}})$ through a cluttered environment. We assume the robotic assistant is feedback controlled such that the commonly manipulated object follows a virtual object impedance that is rendered with inertia matrix $\boldsymbol{M}$, and virtual viscous friction $\boldsymbol{D}$

$$\boldsymbol{M}\ddot{\boldsymbol{x}} + \boldsymbol{D}\dot{\boldsymbol{x}} = \boldsymbol{u}_h + \boldsymbol{u}_r, \tag{1}$$

where $\boldsymbol{x} \in \mathcal{C}_{\text{free}}$ denotes the object pose and $\boldsymbol{u}_h$ and $\boldsymbol{u}_r$ the effective wrenches by human and robot, respectively.

Both cooperation partners have fixed grasp points and are assumed to have their individual plans $\tau_{\boldsymbol{x}(t),h}(s)$ and $\tau_{\boldsymbol{x}(t),r}(s)$ with index $s \in [0, 1]$ that is dependent on the current pose $\boldsymbol{x}(t)$. This implies that the object's pose trajectory $\boldsymbol{x}(t)$ is the result of a continuous plan negotiation through the corresponding input forces $\boldsymbol{u}_h$ and $\boldsymbol{u}_r$ of the human and the robot partner. In the most general case the object pose trajectory deviates from the initially desired paths $\tau_{\boldsymbol{x}(0),h}(s)$ and $\tau_{\boldsymbol{x}(0),r}(s)$ at the starting position $\boldsymbol{x}(0)$ of the human and the robotic partner. The challenge is now to find the appropriate plan $\tau_{\boldsymbol{x}(t),r}$ to generate the robot force contribution, which considers human-preferences as well as environment constraints.

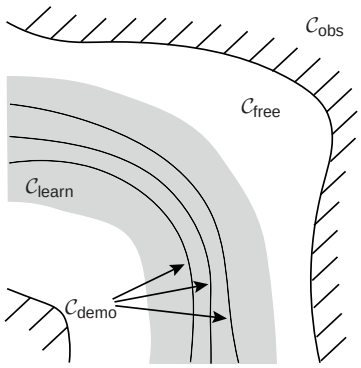


Fig. 2: Subsets of the configuration space $\mathcal{C} = \mathbb{R}^2$

To address this problem, we first investigate a planning-based and a learning-based strategy individually.

In a planning-based approach the robot computes its motion plan based on environmental constraints. In consequence, the robot plan will generally differ from the human plan $\tau_{\boldsymbol{x}(t),r}(s) \neq \tau_{\boldsymbol{x}(t),h}(s)$ and online negotiation is required. As a result the motion of the dyad may deviate from the robot plan $\tau_{\boldsymbol{x}(t),r}(s)$ as determined from the current pose $\boldsymbol{x}(t)$. Feedback motion planning strategies - in contrast to path planning strategies - prepare the goal-directed robotic assistant for this property through a continuous adjustment of the robot's desired path $\tau_{\boldsymbol{x}(t),r}(s)$ from the current configuration to the goal configuration. Furthermore, feedback motion planning suits the need for instant decision making due to the physical coupling. In contrast to path planning, feedback motion planning strategies prepare the goal-directed robotic assistant for this property through a continuous adjustment of the robot's desired path $\tau_{\boldsymbol{x}(t),r}(s)$ from the current configuration $\boldsymbol{x}(t)$ to the goal configuration $\boldsymbol{x}(t_{\text{goal}})$. The robot uses the desired path $\tau_{\boldsymbol{x}(t),r}(s)$ to compute its contribution $\boldsymbol{u}_{\text{plan}}$ as one option for $\boldsymbol{u}_r$.

In a learning-based approach the robot initially behaves *passive* $\boldsymbol{u}_r = 0$ during the task execution and observes the executed trajectory which can be assumed to be close to the human partner's intended path $\tau_{\boldsymbol{x}(0),h}(s)$. In subsequent trials the robot adopts the human preferred path as its own motion plan $\tau_{\boldsymbol{x}(0),r}(s) \leftarrow \tau_{\boldsymbol{x}(0),h}(s)$. Note that the space of learned trajectories is generally only a subspace of the free space, but generally the subspace of learned trajectories is larger than the space of demonstrated trajectories $\mathcal{C}_{\text{demo}} \subseteq \mathcal{C}_{\text{learn}} \subseteq \mathcal{C}_{\text{free}}$, see Fig. 2.

Two representative algorithms from each field, the sampling-based neighborhood graph (SNG) for feedback planning [6] and the time-based HMM (tHMM) learning, recognition and a motion prediction framework [10], [14] are considered in the following and briefly reviewed.

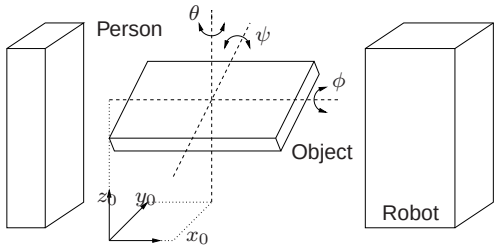


Fig. 3: 6-DoF reduced geometric model of the cooperation partners and the object used for calculation of free configuration space $\mathcal{C}_{\text{free}}$

B. Feedback motion planning

A feedback motion planning scheme generates a feedback function $\boldsymbol{k}(\boldsymbol{x})$ for all positions $\boldsymbol{x} \in \mathcal{C}_{\text{free}}$ in the accessible configuration space $\mathcal{C}_{\text{free}}$. The SNG is a very comprehensible method, sufficiently efficient to cover higher-dimensional configuration spaces applicable for settings with 6-dimensional object poses at a viable resolution as required for our large-scale scenario. A navigation function is precomputed such that goal-directed behavior during execution requires only straightforward lookups. In brief, the method consists of the following steps:

For each new map given by $\mathcal{C}_{\text{free}}$

1) Cover $\mathcal{C}_{\text{free}}$ with n-dimensional hyper balls until a specified percentage a (e.g. .99) is covered with desired probability P_c (e.g. .99).
2) Create a graph representing connectedness of hyper balls.

For each new goal $\boldsymbol{x}_{\text{goal}}$

3) Identify hyper ball containing $\boldsymbol{x}_{\text{goal}}$
4) Generate priority graph e.g. using Dijkstra's algorithm.

During execution

5) Calculate direction to next best hyper ball

Given a desired magnitude of velocity $\dot{x}_d$ and a desired maximum robot force level, the force[1] vector $\boldsymbol{u}_{\text{plan}}$ is calculated straightforwardly using the feedback control law

$$\boldsymbol{u}_{\text{plan}} = \dot{x}_d \frac{\boldsymbol{k}(\boldsymbol{x})}{\|\boldsymbol{k}(\boldsymbol{x})\|} - \dot{\boldsymbol{x}}. \tag{2}$$

Applying the result according to (1) with $\boldsymbol{u}_r = \boldsymbol{u}_{\text{plan}}$ results in the pure planning-based behavior.

The articulated geometric model with six DoF approximating the human, object and robot for a cooperative transport task is depicted in Fig. 3. The object itself can be moved in all six degrees of freedom (DoF). For simplicity in this work we assume that a human partner and a robot control their posture such that they keep a constant configuration relative to the object on the two-dimensional ground plane (x_0,y_0). Aspects of this topic are discussed in [18].

C. Learning from demonstration

Learning-based approaches to active physical assistance aim to imitate the human partner's behavior during demonstrations in order to take over the task effort instead of planning a task-directed behavior. In contrast to planning-based approaches, this can adapt to human preferences.

[1]Force in the 2D case, wrench in the 6D case

The tHMM framework is encodes human trajectory demonstrations efficiently and most importantly constrains motion reproduction to the configuration space $\mathcal{C}_{\text{learn}} \subseteq \mathcal{C}_{\text{free}}$ (i.e. a zone of influence) around the demonstrated configurations $\mathcal{C}_{\text{demo}} \subseteq \mathcal{C}_{\text{learn}}$. We combine Hidden-Markov-Models (HMM) with Gaussian Mixture Regression (GMR) in a probabilistic incremental-learning framework that allows modeling and reproduction of smooth trajectories and is well suited to ensure intuitive motion generation in physical human-robot cooperation. Observed force and motion signals are encoded in time-based hidden-markov models, see [10], [14]. These models are then used to recognize previously observed patterns and the generalized output of the model is used to generate a short-term prediction of the human desired trajectory. To this end the Viterbi algorithm is used for state estimation in time domain. The predicted mean position/velocity setpoint $[\boldsymbol{x}_{\text{pred}}\ \dot{\boldsymbol{x}}_{\text{pred}}]^T$ for time serves as tracking reference in a PD control scheme

$$\boldsymbol{u}_{\text{learn}} = \boldsymbol{K}_p(\boldsymbol{x}_{\text{pred}} - \boldsymbol{x}) + \boldsymbol{K}_d(\dot{\boldsymbol{x}}_{\text{pred}} - \dot{\boldsymbol{x}}), \tag{3}$$

where $\boldsymbol{K}_p$ and $\boldsymbol{K}_d$ denote the proportional and derivative control gains, respectively, to generate the control input $\boldsymbol{u}_{\text{learn}}$. The pure learning-based behavior is synthesized following $\boldsymbol{u}_r = \boldsymbol{u}_{\text{learn}}$. For a detailed explanation of this method, see [10].

III. Synergies of Motion Generation Algorithms

The planning and learning-based approaches described in Section II generate an active contribution behavior for cooperative manipulation tasks. With respect to a number of properties however, these schemes show different and even complementary strengths. In the following, some of these aspects are discussed and a set strategies exploiting these strengths are derived.

A. Properties of motion generation schemes

Most obviously, the learning-based approach is restricted to zones of influence $\mathcal{C}_{\text{learn}}$ around previous observations $\mathcal{C}_{\text{demo}}$. This renders the assistant completely reactive for any first demonstration of a path. In contrast, the feedback motion planning algorithm allows the robotic assistant to contribute anywhere in the entire reachable set $\mathcal{C}_{\text{free}}$. As the planning strategy does not incrementally evolve or adapt to the user over time, any improvement of team performance occurs due to the human partner. In contrast, the incremental learning algorithm tHMM adapts not only to human preferences. Those preferences might even change and adaptation is still provided through task repetition. This assumes that there is a stable mutual adaptation process and path demonstrations remain homotopic (i.e. can be continuously warped from one demonstration to another) which is due to the fact that the environment information is only implicitly contained in the learned models. The problem of furcations in demonstrations is addressed in [10] but excluded from this paper for the sake of simplicity. Nevertheless, the number of design parameters of the learning-based approach is significantly

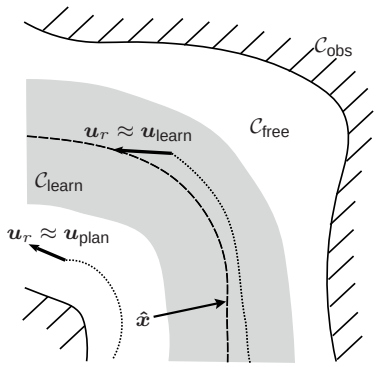


Fig. 4: Region of high confidence around mean generalized output $\hat{\boldsymbol{x}}$ of the learned trajectory model derived from previous demonstrations.

larger and requires more expertise than the parameter space of the planning algorithm. While the SNG approach requires only desired values for precision and coverage, the learning framework requires a desired number of states, number of Gaussians per state and a weighting of input importance. The computational complexity of the the learning framework, however, scales linearly with the input dimension while the SNG free space approximation in planning complexity scales exponentially with the output dimension. However, the high computational effort of free space approximation is necessary only once per scene and can be prepared in advance. The remaining computational effort per new goal configuration is tractable in our envisaged scenarios, the lookup of the control input during execution is negligible. The learning framework requires reasonable computational effort for Expectation Maximization after each observation as well as Viterbi and regression during execution.

B. Synergy strategies

We investigate, how the complementary properties of these two algorithms for motion generation discussed above can be beneficially combined such that synergies arise. Three methods to exploit synergies are derived in the following:

1) Plan-based initialization: The learning algorithm is initialized with a simulated trajectory of the feedback planner. This allows the learning-based approach to render a goal-oriented behavior at the first execution of a specific task. This is supposed to overcome the inability of the learning-based method to assist actively during the first demonstration of a path or sub-path. Therefore, one trajectory from start configuration to goal configuration is simulated with human force $\boldsymbol{u}_h = 0$ and $\boldsymbol{u}_r = \boldsymbol{u}_{\text{plan}}$ according to (1). This artificial observation serves as initial guess.

2) Prediction-quality-based homotopy blending: This strategy exploits the probabilistic characteristics of the motion recognition and prediction algorithm. As recognition uncertainty grows, the goal-oriented fall-back solution generated by the feedback planning algorithm is activated. A homotopy $\mathcal{H}(t, \gamma)$ blends between the two output functions

$$\boldsymbol{u}_{\text{r}} = \gamma \boldsymbol{u}_{\text{learn}} + (1 - \gamma)\boldsymbol{u}_{\text{plan}},$$

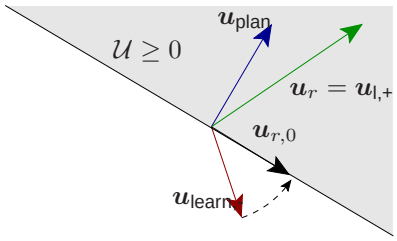


Fig. 5: 2D example of force vectors and solution according to MCDM fusion

with prediction certainty $\gamma \in [0, 1]$. The outputs of the planning algorithm and the learning algorithm are denoted $\boldsymbol{u}_{\text{plan}}$ and $\boldsymbol{u}_{\text{learn}}$ respectively. Similar to [9], we regard the unnormalized likelihood of the state estimate as a measure for certainty γ. This strategy overcomes the limitation of the learning-based algorithm to produce active task contributions only within the zone of influence $\mathcal{C}_{\text{learn}}$ and blends continuously into the feedback motion plan valid for the entire accessible configuration space $\mathcal{C}_{\text{free}}$.

Fig. 4 depicts two examples for the selection of $\boldsymbol{u}_r$ depending on the confidence gained from observations (dotted lines). Trajectories in the region $\mathcal{C}_{\text{learn}}$ around previous demonstrations result in high confidence values and $\boldsymbol{u}_{\text{learn}}$ dominates the control input. Note that for the sake of clarity Fig. 4 depicts only the x and y position components of the example whereas velocity and force signals are equally part of the learning, recognition and prediction algorithm and influence the level of confidence.

3) Cost-based fusion: This method is based on the parallel evaluation of learning and planning-based strategies and a cost-based fusion using a hierarchical multi-criteria decision making (MCDM) algorithm on the efficient (Pareto) frontier. Two different utility criteria are evaluated, one from the planning-based approach and one from the learning-based approach.

The utility function $\mathcal{U}(\boldsymbol{x}, \dot{\boldsymbol{x}}, \boldsymbol{u})$ of the planning approach is calculated through simulation of the effects of different forces $\boldsymbol{u}$ on the *cost to go* by querying the feedback plan. The utility is maximized for force vectors along the planning algorithm output $\boldsymbol{u}_{\text{plan}}$ and and minimized for the opposite direction. Values are normalized to the interval $[-1, 1]$. The value of 0 is cost-neutral.

The utility function of the learning-based algorithm is described by the prediction certainty $\gamma \in [0, 1]$. The certainty value of 1 is reached at the mean direction of the motion prediction according to the unnormalized likelihood of the prediction. Both utility functions are convex.

A hierarchical MCDM-scheme is deployed: The output of the learning-based approach is evaluated for its utility $\mathcal{U}(\boldsymbol{x}, \dot{\boldsymbol{x}}, \boldsymbol{u}_{\text{learn}})$. In case of positive utility, the output of the learning-based approach is accepted, $\boldsymbol{u}_r = \boldsymbol{u}_{\text{learn}}$. Otherwise, the closest direction to $\boldsymbol{u}_{\text{learn}}$ on the efficient frontier with non-negative utility is selected, $\mathcal{U}(\boldsymbol{x}, \dot{\boldsymbol{x}}, \boldsymbol{u}_r) \geq 0$.

$$\boldsymbol{u}_r = \underset{\boldsymbol{u} \in \{\boldsymbol{u}_i | \mathcal{U}(\boldsymbol{x}, \dot{\boldsymbol{x}}, \boldsymbol{u}_i) \geq 0 \cap |\boldsymbol{u}_i| \leq |\boldsymbol{u}_{\text{learn}}|\}}{\arg\max} \boldsymbol{u}^T \boldsymbol{u}_{\text{learn}} \qquad (4)$$

Fig. 5 depicts an exemplary output $\boldsymbol{u}_{\text{plan}}$ of the feedback-

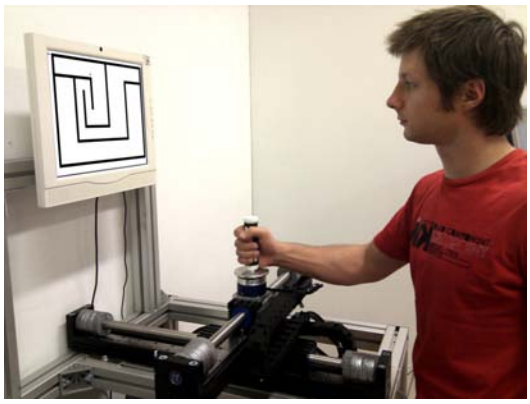


Fig. 6: Subject operating 2-DoF virtual-reality setup

planning algorithm and corresponding directions with non-negative utility $\mathcal{U}(\boldsymbol{x}, \dot{\boldsymbol{x}}, \boldsymbol{u}_r) \geq 0$. Any output by the motion reproduction algorithm $\boldsymbol{u}_{\text{learn},+}$ with non-negative utility is accepted. Outputs of the motion reproduction algorithm $\boldsymbol{u}_{\text{learn},-}$ with negative utility are rotated to the closest direction with non-negative utility.

C. Control architecture

The control algorithm embedding the algorithms proposed above is depicted in Fig. 7. Both methods continuously produce wrenches $\boldsymbol{u}_{\text{plan}}$, $\boldsymbol{u}_{\text{learn}}$ that act on a virtual admittance in superposition with the human partner's force $\boldsymbol{u}_h$ exerted through the object on the end effector of the robot.

IV. Experimental Evaluation

In order to examine the complementary properties of the proposed approaches, two experiments are conducted: a human user study in a two-dimensional virtual scenario is used to evaluate the approaches in terms of physical measures under controlled conditions. Furthermore, we examine the feasibility in a large-scale setting involving kinesthetic interaction between a human and a human-sized mobile robot with manipulators.

A. 2-DoF Virtual Scenario

A small pilot study in a virtual reality scenario was conducted to evaluate the performance of our proposed approaches. Seven non-paid participants (age mean: 27.1, std: 1.5) were asked to move a virtual point mass object of 100 kg, through a simple maze from a starting configuration to a final configuration through the scene without colliding with the virtual obstacles visually and

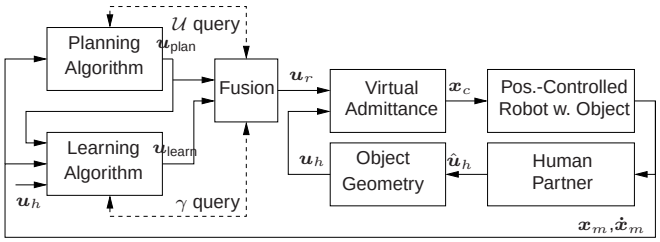
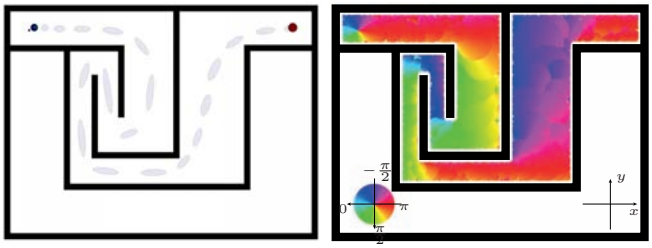


Fig. 7: Control structure



(a) Left-to-right HMM (b) Feedback motion plan

Fig. 8: The utilized motion generation algorithms in 2D

haptically displayed. The virtual object was exposed to a virtual viscous friction of $400\,\text{Ns/m}$. Each participant repeated the experiment five times per each of five conditions after five training trials without assistance: a) assisted by feedback planning, b) assisted by learning-based algorithm, c) learning-based, but pre-initialized with simulated feedback planning, d) prediction-quality-based homotopy blending and e) multi-criteria decision making. The virtual-reality interface consists of a two degrees-of-freedom (anteroposterior and mediolateral plane of the user standing in front) linear-actuated device (*ThrustTube*) with a free-spinning handle (superoinferior direction of the user) at the grasp point. The control algorithm is implemented in *Matlab/Simulink's Real-Time Workshop* and executed on *Linux Preempt/RT* at a sampling rate of 1 kHz. Attached to the handle, a force/torque sensor (*JR3*) measures the human force input. The virtual scene is visually represented on a display placed on top of the interface, see Fig. 6. The displayed task to transport a virtual object is visually represented by a filled red circle and the target position in the upper left corner of the maze (blue dot), see Fig. 8(a). Collisions with the virtual walls should be avoided. For illustration, a two-dimensional feedback plan is depicted in Fig. 8(b) where the hue represents the direction towards the goal which is expressed as constant desired velocity vector. Table I exhibits the constants used to parameterize the experiment.

B. Quantitative measures

We evaluate the following criteria in order to rate the performance of the proposed approaches:

- Mean root-mean-square (RMS) deviation from the participant's path $\tau_{\boldsymbol{x},H5}$ after five trials

$$x_{\text{RMS}} = \sqrt{\frac{1}{T}\int_0^T d(\boldsymbol{x}(t), \tau_{\boldsymbol{x},H5})^2 dt}$$

with distance for point $\boldsymbol{x}(t)$ to path $\tau_{\boldsymbol{x},H5}$

$$d(\boldsymbol{x}(t), \tau_{\boldsymbol{x},H5}) = min_{x_\tau \in \tau_{\boldsymbol{x},H5}} \|\boldsymbol{x}_\tau - \boldsymbol{x}(t)\|.$$

This measure expresses the required amount of path adaptation by the dyad over trials.

Constant	Equation	Value
Simulated object mass $\boldsymbol{M}$	(1)	$100\,\mathrm{kg} \cdot \boldsymbol{I}_2$
Simulated viscous friction $\boldsymbol{D}$	(1)	$400\,\frac{\mathrm{Ns}}{\mathrm{m}} \cdot \boldsymbol{I}_2$
Desired velocity magnitude x_d	(2)	$0.025\,\frac{\mathrm{m}}{\mathrm{s}}$
Proportional gain $\boldsymbol{K}_p$	(3)	$300\,\frac{\mathrm{N}}{\mathrm{m}}$
Derivative gain $\boldsymbol{K}_d$	(3)	$60\,\frac{\mathrm{Ns}}{\mathrm{m}}$

TABLE I: Control parameters used in 2-DoF experiment

- Mean disagreement $\boldsymbol{u_D}$ which is defined orientation invariant:

$$\boldsymbol{u}_D = \begin{cases} \dfrac{-\boldsymbol{u}_h}{\|\boldsymbol{u}_h\|} \cdot \boldsymbol{u}_r, & \text{if } \boldsymbol{u}_h \cdot \boldsymbol{u}_r < 0 \wedge\ \boldsymbol{u}_h \neq \boldsymbol{0} \\ 0, & \text{otherwise.} \end{cases}$$

- Mean completion time T_{mean}.

C. Quantitative results from human user study

The evaluation results of the properties of the planning-based and learning-based approaches in terms of physical measures are depicted in Fig. 9.

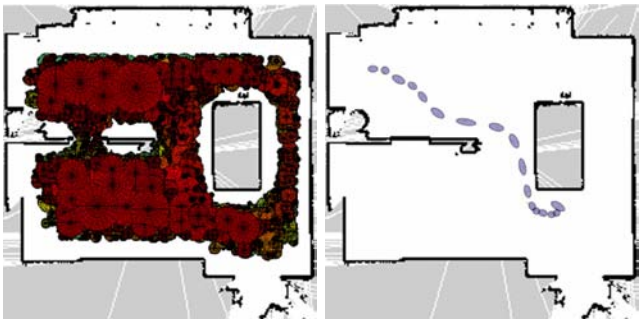
Regarding the RMS deviation over trials, it is visible that the conditions d) Homotopy blending and e) MCDM require the least adaptation and are in a similar range as the b), the pure learning-based approach. Condition a), the planning-based approach leads to stronger adaptation of the human. The measurements of mean disagreement show that the conditions a), d) and e) perform equally well. The mean disagreement of these conditions is lower than in the pure learning-based condition. The completion times of the fusion strategies d) and e) are similar to those of the planning-based approach and lower than in the learning-based strategy. Condition c), the plan-initialized learning method performs better than the pure learning-based approach in the first trial.

From the quantitative results it is visible that the fusion strategies d) and e) combine the strengths of the planning-based and the learning-based approaches a) and b). These fusion methods outperform the planning-based approach regarding the required adaptation of the human and leads to small disagreement compared to the learning-based approach. The completion time under these conditions is on the level of the planning approach throughout all trials and is significantly lower than the unassisted condition in the first trial of the learning-based approach.

A pre-initialization of the learning algorithm is simple to implement but trains the trajectory models according to the robotic desired solution which deviates from the human solution. This deviation is uncomfortable for the human user and vanishes only after a larger number of repetitions.

V. Experiments in 6 Degrees of Freedom

Challenges arise when the proposed methods are applied to a cooperative load transport scenario. Properties such as scalability to higher degrees of freedom, robustness to sources of noise in autonomous mobile manipulation and safe behavior outside the expected motion corridor are evaluated in a proof-of-concept implementation on our highly integrated experimental system that has been instrumental to validate different aspects of human-robot collaboration[2].



(a) Feedback motion plan (b) Left-to-right HMM

Fig. 10: Motion generation algorithms in 6D (x_0/y_0 components)

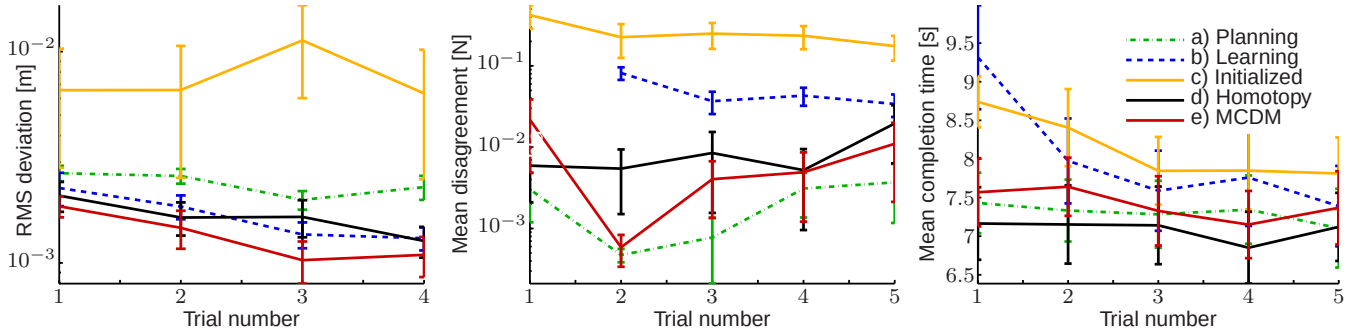
The mobile robot used in this experiment (see Fig. 1) locomotes with its four-wheeled omni-directional mobile platform which offers roughly human-like maneuverability and smooth motion [19]. Two identical anthropomorphic backlash-free 7-degrees-of-freedom manipulators are front-mounted at the top of the main chassis to provide a human-like working space [20]. Mounted onto *JR3* wrench sensors, the manipulators are equipped with *Schunk PG70* two-finger parallel grippers that allow a tight grasp of the object. Lithium-ion polymer batteries power the system for long periods without recharging. A point cloud of the environment is acquired using a tilted *Hokuyo UBG* laser range finder. Two *Sick S300* laser range finders scan for obstacles above the surrounding floor. The cooperatively manipulated objects are a $1\,\mathrm{m} \times 0.5\,\mathrm{m}$ Styrofoam board and a Mini's $1.1\,\mathrm{m}$ long steel bumper. They are moved by the human-robot dyad through the cluttered $10\,\mathrm{m} \times 10\,\mathrm{m}$ laboratory environment. A protoypical path is depicted in Fig. 10(b).

The software framework used in this experiment is based on our *ARCADE* framework [21] suited for rapid prototyping of perception-cognition-action loops in complex human-robot teams scenarios. *ROS* is utilized for self localization. The admittance-type control scheme is implemented in *Matlab/Simulink's Real-Time Workshop* and is executed on *Linux Preempt/RT* at a sampling rate of $1\,\mathrm{kHz}$.

A. Results from experiments in 6D

For fast computation, parallelized implementations on an *Intel Core i7 920 at 2.67GHz* are utilized. The computation times for our prototypical 6D-problem are given in Table II. A projected feedback plan is depicted in Fig. 10(a). Our proof-of-concept implementation shows the feasibility of the approach for large-scale human-robot setups in the domain of cooperative load transport. The evaluation of different start and goal configurations shows how the planning algorithm suggests to tilt the transported Styrofoam board in order to pass through a narrow door.

[2]see `http://www.cotesys.org/newsroom/videos`



(a) RMS path deviation (b) Disagreement (c) Completion time

Fig. 9: Evolution of quantitative parameters over trials

Computation step	SNG	tHMM
Calculation of neighborhood graph from point cloud ($\alpha = P_c = 0.985$)	639 s	
Planning per new goal configuration	< 1 s	
Expectation Maximization after each observation		~ 5 s
Direction lookup per control iteration	$< 50\,\mu$s	
Viterbi and regression during execution		$< 500\,\mu$s

TABLE II: Computation times of SNG and tHMM in 6D

VI. Conclusion

In this paper we propose a novel approach to goal-directed behavior in physical human-robot interaction (pHRI) which is based on the synergetic combination of planning and learning-based approaches. In particular, a feedback planning scheme is suggested for pHRI. A qualitative comparison with learning-based approaches for the same application unveils opportunities for fusion of both approaches to overcome limitations of the individual algorithms. From our experiments in two and six degrees of freedom, we can summarize a number of crucial properties relevant to the application of cooperative load transport. Experimental results show that both investigated fundamental strategies, a learning-based as well as a planning-based approach yield feasible implementations of active robotic assistants for lifelike tasks in six degrees of freedom. From our quantitative evaluation, we can state that both approaches can benefit from each other, given a fusion algorithm that blends or switches between the two approaches depending on the situation.

Acknowledgements

This work is supported in part within the DFG excellence initiative research cluster Cognition for Technical Systems - CoTeSys (`www.cotesys.org`).

References

[1] K. Kosuge, H. Yoshida, and T. Fukuda, "Dynamic control for robot-human collaboration," in *Proc. IEEE Workshop on Robot and Human Communication*, 1993, pp. 398–401.

[2] M. Taïx and D. Flavigné, "Motion Planning with interactive devices," in *Proc. IEEE ECMS*, 2011, pp. 1–6.

[3] J. Rosell, C. Vázquez, A. Pérez, and P. Iñiguez, "Motion Planning for Haptic Guidance," *J. Intell. Robotic. Syst.*, vol. 53, no. 3, pp. 223–245, 2008.

[4] L. Rosenberg, "Virtual Fixtures: Perceptual tools for telerobotic manipulation," in *Proc. IEEE VR*, 1993, pp. 76–82.

[5] C. Vázquez, J. Rosell, L. Chirinos, and O. Domínguez, "Haptic Primitives Guidance Based on the Kautham Path Planner," in *Proc. IEEE/RSJ IROS*, 2010, pp. 4686–4691.

[6] L. Yang and S. LaValle, "The Sampling-Based Neighborhood Graph: An Approach to Computing and Executing Feedback Motion Strategies," *IEEE Trans. Robot. Automat.*, vol. 20, no. 3, pp. 419–432, 2004.

[7] P. Evrard, E. Gribovskaya, S. Calinon, A. Billard, and A. Kheddar, "Teaching Physical Collaborative Tasks: Object-Lifting Case Study with a Humanoid," in *Proc. IEEE Humanoids*, 2009, pp. 399–404.

[8] E. Gribovskaya, A. Kheddar, and A. Billard, "Motion learning and adaptive impedance for robot control during physical interaction with humans," in *Proc. IEEE ICRA*, 2011, pp. 4326–4332.

[9] A. Thobbi, Y. Gu, and W. Sheng, "Using Human Motion Estimation for Human-Robot Cooperative Manipulation," in *Proc. IEEE/RSJ IROS*, 2011, pp. 2873–2878.

[10] J. Medina, M. Lawitzky, A. Mörtl, D. Lee, and S. Hirche, "An Experience-Driven Robotic Assistant Acquiring Human Knowledge to Improve Haptic Cooperation," in *Proc. IEEE/RSJ IROS*, 2011, pp. 2416–2422.

[11] V. Fernandez, C. Balaguer, D. Blanco, and M. Salichs, "Active Human-Mobile Manipulator Cooperation Through Intention Recognition," in *Proc. IEEE ICRA*, 2001, pp. 2668–2673.

[12] N. Jarrasse, J. Paik, V. Pasqui, and G. Morel, "How can human motion prediction increase transparency?" in *Proc. IEEE ICRA*, 2008, pp. 2134–2139.

[13] A. Billard, S. Calinon, R. Dillmann, and S. Schaal, "Robot Programming by Demonstration," in *Handbook of Robotics*, B. Siciliano and O. Khatib, Eds. Springer, 2008, pp. 1371–1394.

[14] D. Lee and C. Ott, "Incremental kinesthetic teaching of motion primitives using the motion refinement tube," *Auton. Robot.*, pp. 1–17, 2011.

[15] M. Lawitzky, A. Mörtl, and S. Hirche, "Load Sharing in Human-Robot Cooperative Manipulation," in *Proc. IEEE Ro-Man*, 2010, pp. 185–191.

[16] A. Mörtl, M. Lawitzky, A. Küçükyılmaz, M. Sezgin, C. Basdogan, and S. Hirche, "The Role of Roles: Physical Cooperation between Humans and Robots," *Int. J. Robot. Res.*, 2012 (to appear).

[17] A. Kirsch, T. Kruse, E. Sisbot, R. Alami, M. Lawitzky, D. Brščić, S. Hirche, P. Basili, and S. Glasauer, "Plan-Based Control of Joint Human-Robot Activities," *Künstl. Intell.*, vol. 24, no. 3, pp. 223–231, 2010.

[18] C. Esteves, G. Archevaleta, and J. Laumond, "Motion Planning for Human-Robot Interaction in Manipulation Tasks," in *Proc. IEEE ICMA*, 2005, pp. 1766–1771.

[19] U. Hanebeck, N. Saldic, and G. Schmidt, "A Modular Wheel System for Mobile Robot Applications," in *Proc. IEEE/RSJ IROS*, 1999, pp. 17–22.

[20] B. Stanczyk and M. Buss, "Development of a Telerobotic System for Exploration of Hazardous Environments," in *Proc. IEEE/RSJ IROS*, 2004, pp. 2532–2537.

[21] D. Althoff, O. Kourakos, M. Lawitzky, A. Mörtl, M. Rambow, F. Rohrmüller, D. Brščić, D. Wollherr, S. Hirche, and M. Buss, "An Architecture for Real-time Control in Multi-robot Systems," in *Cognitive Systems Monog.: Human Centered Robot Systems*, 2009, pp. 43–52.