Workspace Analysis for Planning Human-Robot Interaction Tasks

Nikolaus Vahrenkamp[1], Harry Arnst[1], Mirko Wächter[1], David Schiebener[1],
Panagiotis Sotiropoulos[2], Michal Kowalik[2], and Tamim Asfour[1]

***Abstract*— We present an approach for determining suitable locations for human-robot interaction tasks. Therefore, we introduce the task specific Interaction Workspace as a representation of the workspace that can be accessed by both agents, i.e. the robot and the human. We show how the Interaction Workspace can be efficiently determined for a specific situation by making use of precomputed workspace representations of robot and human. By considering several quality measures related to dexterity and comfort, the Interaction Workspace provides valuable information about potential targets for human robot interaction (e.g. for object hand-over tasks). We evaluate the online performance of building appropriate data structures and show how the approach can be applied in a realistic hand-over use case with the humanoid robot ARMAR-III.**

I. INTRODUCTION

Robots which collaborate with humans in shared environments need to be equipped with planning components that consider the capabilities of the human. Therefore, a representation of the current state of the human together with an estimation of the human's capabilities is essential. The state of the human is usually represented by his or her position and configuration, while the representation of the capabilities is task dependent.

In this work, we are investigating human-robot interaction processes, such as hand-over tasks, and how a beneficial representation of the human capabilities in terms of reaching and comfort can be used to support the autonomous planning of hand-over locations. Instead of using heuristics, which can be used to solve simple setups (e.g. the robot is located directly in front of the human), our approach allows for fully autonomous computation of suitable hand-over poses, even in complex and constrained environments. The approach incorporates a pre-computed representation of the reachable workspace of the human and uses this information in an online manner to determine the shared Interaction Workspace of robot and human. The Interaction Workspace can take into account online information, e.g. the effort that is needed by the human to reach a specific location from his or her current state, which allows to search for optimal interaction targets. Due to the efficient generation of the Interaction Workspace, the approach is suitable for planning online human-robot interaction tasks, such as handing over objects from the robot to a human collaborator. We evaluate the approach in simulation as well as in a maintenance scenario in which a human technician is assisted by the humanoid robot ARMAR-III [1].

[1]High Performance Humanoid Technologies Lab, Institute for Anthropomatics and Robotics, Karlsruhe Institute of Technology (KIT), Germany

[2]Robotics Research Team at Ocado Technology, Hatfield, UK

The research leading to these results has received funding from the European Unions Horizon 2020 Research and Innovation programme under grant agreement No 643950 (SecondHands).

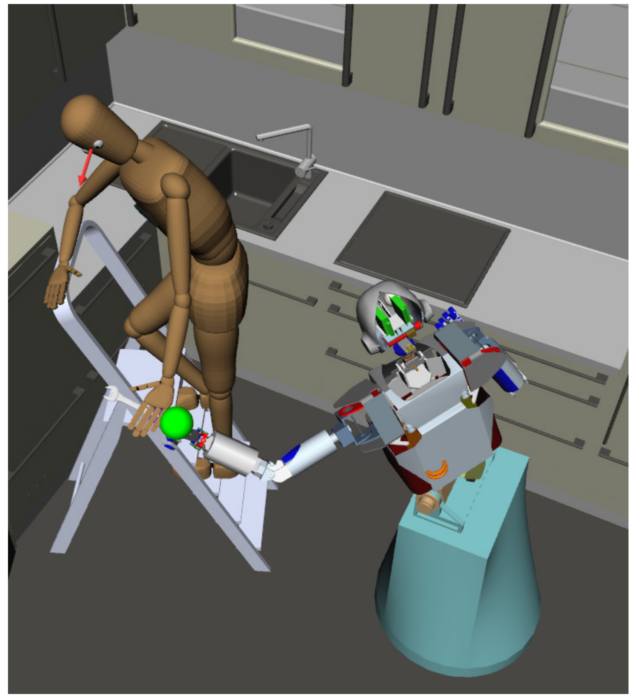

Fig. 1. ARMAR-III passing an object to a human in simulation. The hand-over target pose is determined by querying the Interaction Workspace which is computed by intersecting the workspaces of robot and human.

II. RELATED WORK

A. Work Space Analysis in Robotics

Analyzing the workspace of a manipulator is a useful step for many applications in the context of robotic manipulation. Especially for tasks like grasp and motion planning it is important to have a representation of the manipulator's capabilities within specific regions in Cartesian space. These capabilities can be represented by reachability information [2] or more sophisticated quality distributions defined in workspace.

Zacharias et al. [3] introduces a workspace representation that stores directional structure information to describe a manipulator's capability to reach certain workspace regions from different directions.

In earlier work, we proposed a 6D voxelized data structure to represent the workspace of a manipulator [4]. Each voxel holds a manipulability reference value describing the manipulator's local dexterity.

Porges et al. [5] analyzes different approaches to create such 6D voxelized data structures. Furthermore an overview of different methods for evaluating voxel quality is given.

B. Representation of the Human Workspace

Performing human-aware interactions is a key ability for robots that have to act in human-centered environments. Especially in the context of planning everyday interactions like handover tasks there exist several works that analyze the human workspace to account for various human-oriented constraints and preferences such as comfort and safety. In order to satisfy human-oriented constraints (e.g. comfort of arm postures) one can define quantitative quality measures expressing the quality of e.g. arm postures regarding respective constraints.

This can be done by using a cost map approach, as proposed in [6], [7]. In these works, the human workspace is represented by a 3D grid and analyzed based on various cost functions regarding human's comfort, safety and field of view. Grid cells with minimal costs are assumed to be potentially good handover positions. Robot motions are planned via a path search in human's workspace [6] or robot's configuration space [7]. Paths that minimize the overall cost of the cells they contain, should result in potentially safe and legible robot motions.

Other works use an intuitive approach to account for human-oriented constraints [8], [9], [10]. In these works, the human workspace is represented by a set of 3D grids which are referred to as "mightability maps". Each mightability map stores binary reachability and visibility information of the human. Candidate handover positions are determined through set operations on the robot's and human's mightability maps.

One common point across the aforementioned related work is that workspaces are represented by 3D grids. For planning handover tasks with smaller objects, orientations can be disregarded, hence this approach is sufficient. However, in general the position where the human grasps an object depends heavily from its orientation and therefore from the orientation of the robot's hand holding the object. To consider not only positions, but also orientations, we use in this paper 6D voxelized data structures (as described in [4]) to represent the workspace of the human and the robot.

III. APPROACH

An overview of the proposed approach for representing the human-robot Interaction Workspace is depicted in Fig. 2. The workspace of the human is created for the human reference model of the Master Motor Map (MMM) framework [11], [12]. Since the MMM framework is capable to adapting the size of the MMM model, this representation can either be created for a formerly known subject height, or the model together with the workspace representation can be adapted online according to an estimated subject size. In addition,

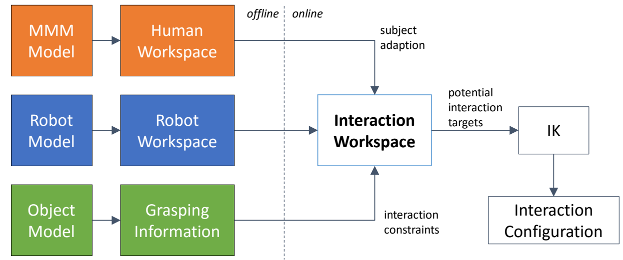

Fig. 2. The Interaction Workspace takes into account the human's and the robot's workspace representation together with additional task constraints. Based on this data structure, potential interaction targets can be generated which have to be verified by an inverse kinematics solver. The resulting interaction configuration can be executed by the robot.

the robot's workspace and grasping information are created during the offline step with the approaches of [4] and [13].

During online processing, the information is used to create an *Interaction Workspace* that can be queried for potential interaction targets. The IK solver finally creates an *Interaction Configuration*.

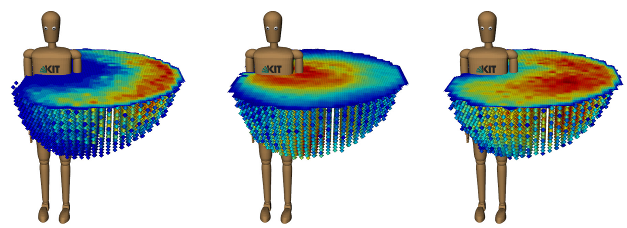

Fig. 3. Visualization of a cut through the 6D workspaces of the human model's left arm (7 DoF). (a) Quality distribution regarding local dexterity of the human arm (red: high dexterity, blue: low dexterity). (b) Quality distribution regarding energy consumption of arm postures (red: low effort, blue: high effort). (c) Exemplary combination of both workspace representations.

A. Human Workspace Analysis

In this work the capabilities of the human are represented by a grid of voxels in 6D pose space, where a pose is described by (t, R), with $t \in \mathbb{R}^3$ the translational and $R \in SO(3)$ the rotational part. Each voxel is associated with a quality value $q \in [0, 1]$ indicating the human arm's local dexterity. Determining the human arm's local dexterity is based on manipulability anaysis described in [14]. Fig. 3 (a) shows a quality distribution representing the local dexterity of the human left arm (7 DoF), whereas the different quality values are encoded by color.

For the offline computation of the human's workspace we use a sampling-based approach. A large set of random arm configurations is generated, for every configuration the pose of the Tool Center Point (TCP) is computed via forward kinematics and the voxel that corresponds to the TCP pose is evaluated regarding the dexterity of the human arm. Since the workspace grid is linked to the human model's shoulder, it adjusts its position and orientation when the human is moving (e.g. bending over or crouching).

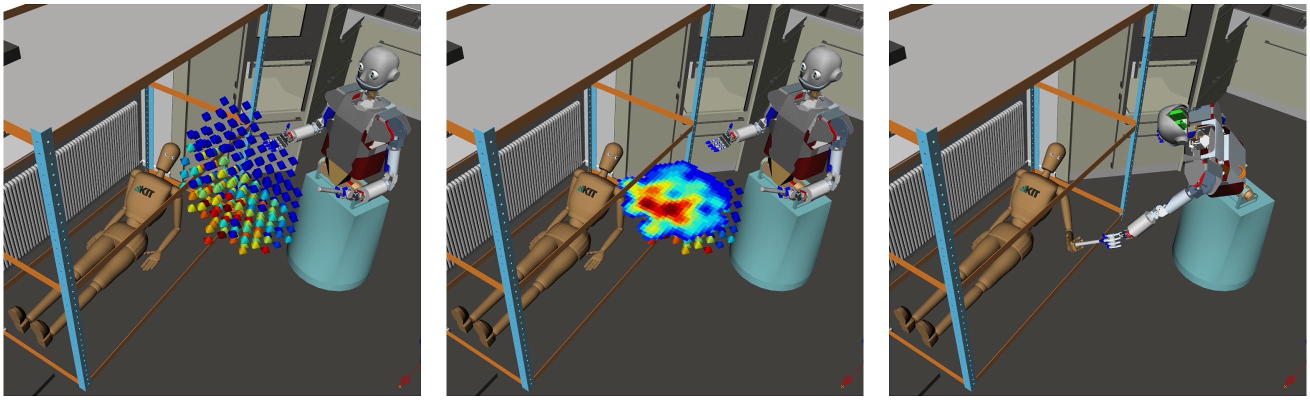

Fig. 4. The robot is supposed to hand-over a mallet to a person lying on the ground. (a) Interaction Workspace. (b) Cut through the Interaction Workspace. (c) An interaction configuration.

B. An Online Quality Measure for Human Postures

An offline analysis of the human's capabilities (Sec. III-A) provides valuable information for planning interaction tasks. However, since Human-Robot Interaction (especially in the context of object hand-over) is a highly dynamic process, we have to consider the human's current state to account for human-oriented constraints like comfort and safety. Therefore, we introduce a quality measure to evaluate potential interaction targets in an online manner considering the human model's current state.

For a given potential interaction target (i.e. a 6D pose) we obtain the corresponding target configuration θ_{target} of the human arm via inverse kinematics. Furthermore, we capture the current gaze direction and posture of the human model to account for visibility and safety constraints.

In the following, we introduce several costs functions, which are combined to evaluate a quality value for a given interaction target in workspace.

The first two cost functions are related to the *travel cost* and *spatial error* of Rosenbaum et al.'s posture-based motion planning model [15]. However, we do not restrict potential target poses to be on the sagittal or horizontal plane only and we assume optimal movement times for each joint of the human arm. Potential interaction targets, that minimize the angular and spatial displacement costs, are assumed to be promising in terms of human arm comfort. Besides human arm comfort we also consider the human's current gaze direction, safety constraints (i.e. collision free interactions) and the effort of human arm postures.

a) Joint space travel: For a given target configuration of the human arm a cost based on the angular displacement of all joints is defined by

$$c_{joint} = \sum_{i=1}^{n} \alpha_i \cdot (\theta_{current,i} - \theta_{target,i})^2 \qquad (1)$$

where $\alpha_i \in \mathbb{R}^+$ is a weight applied to the ith joint, $\theta_{current,i}$ is the current angle and $\theta_{target,i}$ is the target angle of the ith joint.

b) Work space travel: For a given interaction target, a cost based on the spatial displacement of the human arm's TCP is defined by

$$c_{spatial} = \|p_{current} - p_{target}\|_2 \qquad (2)$$

where $p_{current} \in \mathbb{R}^3$ is the current position and $p_{target} \in \mathbb{R}^3$ is the target position of the human arm's TCP.

c) Visibility: The visibility cost $c_{visibility} \in [0, \infty]$ is based on the angle between the human's current gaze direction and the given potential interaction target. The greater the angle, the higher the cost. The human's gaze direction is approximated by a vector shown in Fig. 1 (red arrow).

If the gaze of the human cannot be estimated by the robot's perception system, this quality value can either be omitted, or a standard gaze direction can be assumed which e.g. can be derived from the orientation of the upper body.

d) Collision: The safety cost $c_{safety} \in [0, \infty]$ is based on the distance between the target TCP position and obstacles (e.g. own body). The shorter the distance, the higher the cost. The cost evaluates to infinity on collision and to zero after a maximal distance.

e) Effort: The effort cost $c_{effort} \in [0, \infty]$ is based on the energy consumption of human arm postures. The higher the effort, the higher the cost. Determining the effort of arm postures is accomplished by analyzing the overall torque of all joints of the human model's arm. Energy consumption increases as the overall torque of all joints increases. Arm postures that minimize the total torque of all joints are assumed to be good candidates for comfortable arm postures. This point of view of relating energy consumption of arm postures to human comfort is based on Katayama's *minimum joint torque index* [16]. Fig. 3 (b) shows a quality distribution representing the energy consumption of arm postures for a standing human model.

f) Quality representation: For further processing, we transfer the costs to a quality representation in $[0, 1]$:

$$q_i = 1 - tanh(\beta_i c_i), \qquad (3)$$

with $i \in \{joint, spatial, visibility, safety, effort\}$ and a scaling constant β_i.

g) Total quality: The online quality measure $q_{online} \in [0, 1]$ considering human comfort, visibility, safety and effort of arm postures is defined as a weighted sum of above qualities:

$$q_{online} = \sum_i \omega_i q_i, \tag{4}$$

with $i \in \{joint, spatial, visibility, safety, effort\}$ and the weights $\omega_i \in [0, 1]$ sum up to 1. In our experiments we uniformly distribute the influence of the different qualities by choosing $\omega_i = 1/5$.

The overall quality $q_{total} \in [0, 1]$ is then obtained by a weighted sum of q_{online} and the offline computed local dexterity value of the target arm posture (Sec. III-A).

For planning human-robot interaction tasks, maximizing the overall quality q_{total} results in interaction targets, that are as comfortable, visible and safe as possible.

C. Interaction Workspace

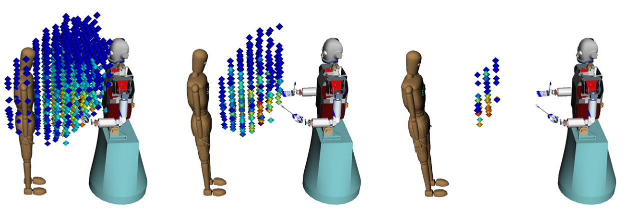

Fig. 5. Interaction Workspace for various distances between robot and human (70cm, 110cm, 150cm). Average number of potential interaction targets and computation times are shown in Table I.

In order to hand over an object to a human, the robot has to determine potential hand-over poses (i.e. potential interaction targets). A potential hand-over pose must meet the following requirements:

- It must be reachable for the robot's TCP.
- It must ensure, that when attained by the robot's TCP, the held object becomes reachable and graspable for the human.
- It must satisfy task dependent constraints.

The set of all potential hand-over poses between the robot and human is represented by a voxel grid in 6D pose space, which will be referred to as *Interaction Workspace*. Since all potential hand-over poses are implicitly reachable for the robot's TCP, the Interaction Workspace is fully covered by the robot's workspace.

The Interaction Workspace directly depends on the actual relation of robot and human and hence, it cannot be created in an offline step. Nevertheless, the workspace representations of robot and human, which were created offline, can be used to efficiently build the Interaction Workspace as an intersection of both workspaces. In addition, quality considerations in terms of dexterity and comfort are taken into account.

As shown in Algorithm 1, the Interaction Workspace W_I is constructed by considering the robot grasp g_{robot}, which defines the transformation form the robot's TCP to the object, and a set of grasps G_{human} that define possible grasps which can be realized by the human in order to receive the object. The entries of W_I are created by going through all grid cells and, in case the corresponding pose is reachable by the robot, updating the entry with the corresponding quality value q. The quality of the current pose is determined on the fly by solving the inverse kinematics for the human and computing the overall quality q_{total} (Sec. III-B). The quality value q_{total} is then stored in the corresponding entry of the Interaction Workspace.

A visualization of the Interaction Workspace for an exemplary scene is shown in Fig. 4 (a) and (b). The human is laying on the ground and the robot is supposed to hand-over an object to the person. The 6D Interaction Workspace W_I is created as described above and visualized in 3D. Fig. 4 (b) shows a cut through W_I, in which the color indicates the quality over all orientational dimensions of W_I.

Algorithm 1 Buildup of the Interaction Workspace

function INTERACTIONWORKSPACE(G_{human}, g_{robot})
  $W_{robot} \leftarrow$ workspace of the robot
  $W_{human} \leftarrow$ workspace of the human
  $W_I \leftarrow$ empty Interaction Workspace
  for all Pose $p_{voxel} \in W_I$ **do**
    if W_{robot}.isReachable(p_{voxel}) **then**
      for all Grasp $g_{human} \in G_{human}$ **do**
        Pose $p_{object} \leftarrow g_{robot} \cdot p_{voxel}$
        Pose $p_{human} \leftarrow p_{object} \cdot g_{human}$
        Configuration $c_{human} \leftarrow$ IK(p_{human})
        Quality $q \leftarrow$quality(c_{human})
        W_I.updateEntry(p_{voxel},q)
  return W_I

D. Determining Suitable Interaction Targets

The Interaction Workspace W_I can be used to identify potential workspace poses at which human-robot interaction can take place. Hence, possible targets for hand-over tasks can be determined in workspace, but they need to be verified by an IK solver, since the workspace representations are approximated and potential collisions with the environment need to be taken into account. As shown in Algorithm 2, W_I is queried for promising poses until an IK solution could be found. In case the IK search fails, the corresponding entry in W_I is removed. A potential interaction target is shown in Fig. 4 (c). The IK query was solved for robot and human.

IV. EVALUATION AND APPLICATION

A. Evaluation

In the following, we evaluate the computation of the Interaction Workspace in a scenario where the robot is placed in front of the human at various distances. The task is to find a suitable configuration for the robot to hand over a wrench that is held in the robot's left hand. The setup involves a 43-DoF model of the humanoid robot ARMAR-III [1] and a 104-DoF model of an average sized human [11], [12]. The

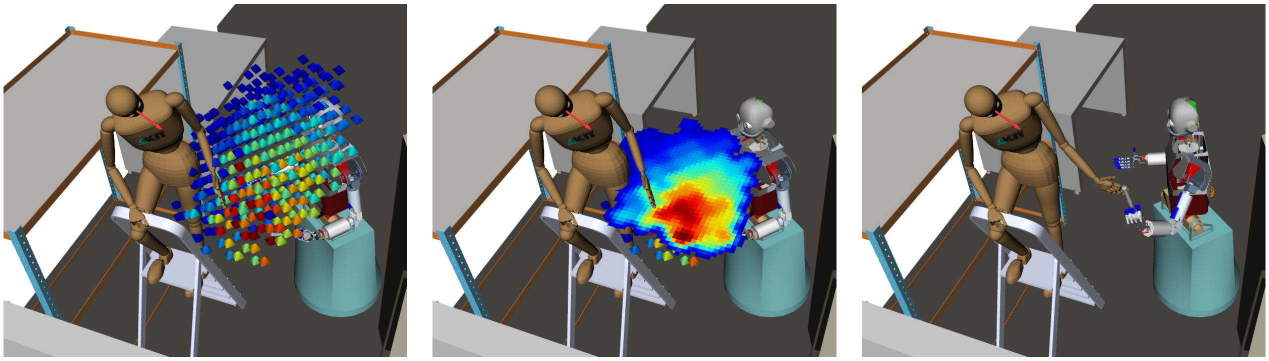

Fig. 6. Scenario of a task to hand over a mallet that is held in the robot's left hand. (a) Interaction Workspace. (b) Cut through the Interaction Workspace. (c) Interaction Target.

Algorithm 2 Compute a Hand-over Target

function SEARCH-HAND-OVER-CONFIGURATION
    $W_I \leftarrow$ Interaction Workspace
    while !Timeout() **do**
        Pose $p \leftarrow$ pose with highest quality in W_I
        Configuration $c \leftarrow$ IK(p)
        if isValid(c) **then**
            return c
        else
            W_I.remove(p)
    return $failure$

Distance (cm)	Potential Targets	Time (s)
70	3576	2.87
90	2290	2.21
110	1110	1.24
130	373	0.51
150	78	0.18

TABLE I
AVERAGE NUMBER OF POTENTIAL INTERACTION TARGETS AND COMPUTATION TIME OF THE INTERACTION WORKSPACE FOR VARIOUS DISTANCES BETWEEN HUMAN AND ROBOT.

simulation is carried out with the C++ robotics simulation package Simox [13]. All tests have been performed on a 2.4 GHz Core i7 PC within a mulit-threaded application.

Table I shows the average number of potential interaction targets and the computation time of the Interaction Workspace for various distances between human and robot. For computing the Interaction Workspace, the human's left arm (7 DoF), as well as the robot's hip and left arm (3 + 7 DoF) were considered. The results show that the computation time and size of the Interaction Workspace is inversely proportional to the distance between human and robot. The greater the distance, the less possibilities do exist to hand over an object to the human, hence the decreasing size of the Interaction Workspace. Since determining a human-oriented quality index for every voxel (i. e. possible interaction target) involves solving the inverse kinematics, the computation time directly depends on the size of the Interaction Workspace (i. e. the number of voxels). Fig. 5 shows the Interaction Workspaces that correspond to the distances of 70cm, 110cm and 150cm.

B. Application to a hand-over task

In this use case, we show how the proposed Interaction Workspace can be used in a realistic setup with the humanoid robot ARMAR-III. The task of the robot is to support a human technician who is performing maintenance works while standing on a ladder.

During the maintenance work, the human operator will from time to time need different tools, which he requests from the robot. The robot has to localize them, grasp them and hand them over to the human. Similarly, the human sometimes hands back objects to the robot that it has to place in an appropriate location.

The whole task is programmed using the statechart-based robot software framework ArmarX [17]. Grasping and placing of the objects is realized using position-based visual servoing [18], and the robot automatically chooses appropriate view directions based on the required objects and its uncertainty about their locations [19]. The ArmarX statechart is visualized in Fig. 7.

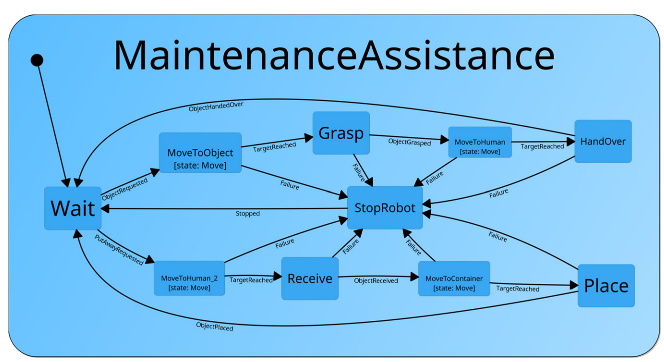

Fig. 7. The ArmarX statechart for the maintenance task.

The handover from robot to human is planned based on the Interaction Workspace as described in Section III. The

Fig. 8. A hand over task executed with ARMAR-III.

resulting data structures are visualized in Fig. 6. The robot hands over the object and waits for the human to grasp it, which is detected using a force-torque sensor in the wrist. Only when the human exerts force on the object by grasping it, the robot releases it (see Fig. 8).

For the handover in the other direction, i.e. from human to robot, the robot has to react to the human's actions and thus can not plan its motion beforehand. Instead, once the human hand comes into reach, the robot moves its hand towards it until a force contact indicates that the object is within its hand, and grasps it.

V. CONCLUSION

We introduced the Interaction Workspace as a representation of the workspace that is accessible by multiple agents in human-robot interaction tasks. The Interaction Workspace provides task-specific information about dexterity and comfort and can be efficiently queried to determine potential targets for human-robot interaction tasks (e.g. object hand-over). It can be constructed in an efficient way by making use of pre-computed workspace representations of robot and human. Due to the possibility to incorporate different quality measures, task-specific constraints can be considered during online processing. We evaluated the performance of the buildup of the Interaction Workspace in different setups. In addition, we showed how the proposed approach can be used for planning hand-over poses in a human-robot interaction scene.

In future work, we plan to apply the Interaction Workspace approach for different applications, e.g. for planning joint actions of human and robot. Additionally, we will investigate if other quality measures, e.g. from biomechanical studies, can be used to improve the representation of the human comfort.

REFERENCES

[1] T. Asfour, P. Azad, N. Vahrenkamp, K. Regenstein, A. Bierbaum, K. Welke, J. Schroeder, and R. Dillmann, "Toward humanoid manipulation in human-centred environments," *Robotics and Autonomous Systems*, vol. 56, no. 1, pp. 54–65, 2008.

[2] N. Vahrenkamp, D. Berenson, T. Asfour, J. Kuffner, and R. Dillmann, "Humanoid motion planning for dual-arm manipulation and re-grasping tasks," in *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2009, pp. 2464–2470.

[3] F. Zacharias, C. Borst, and G. Hirzinger, "Capturing robot workspace structure: representing robot capabilities," in *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2007, pp. 3229–3236.

[4] N. Vahrenkamp, T. Asfour, G. Metta, G. Sandini, and R. Dillmann, "Manipulability analysis," in *2012 12th IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, 2012, pp. 568–573.

[5] O. Porges, T. Stouraitis, C. Borst, and M. A. Roa, "Reachability and capability analysis for manipulation tasks," in *ROBOT2013: First Iberian Robotics Conference*. Springer, 2014, pp. 703–718.

[6] E. A. Sisbot, L. F. Marin, and R. Alami, "Spatial reasoning for human robot interaction," in *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2007, pp. 2281–2287.

[7] J. Mainprice, E. A. Sisbot, L. Jaillet, J. Cortés, R. Alami, and T. Siméon, "Planning human-aware motions using a sampling-based costmap planner," in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, 2011, pp. 5012–5017.

[8] A. K. Pandey and R. Alami, "Mightability maps: A perceptual level decisional framework for co-operative and competitive human-robot interaction," in *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, 2010, pp. 5842–5848.

[9] A. K. Pandey, J.-P. Saut, D. Sidobre, and R. Alami, "Towards planning human-robot interactive manipulation tasks: Task dependent and human oriented autonomous selection of grasp and placement," in *2012 4th IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob)*, 2012, pp. 1371–1376.

[10] A. K. Pandey and R. Alami, "Taskability graph: Towards analyzing effort based agent-agent affordances," in *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*, 2012, pp. 791–796.

[11] P. Azad, T. Asfour, and R. Dillmann, "Toward an unified representation for imitation of human motion on humanoids," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2007, pp. 2558–2563.

[12] O. Terlemez, S. Ulbrich, C. Mandery, M. Do, N. Vahrenkamp, and T. Asfour, "Master motor map (MMM) - framework and toolkit for capturing, representing, and reproducing human motion on humanoid robots," in *IEEE/RAS International Conference on Humanoid Robots (Humanoids)*, 2014, pp. 894–901.

[13] N. Vahrenkamp, M. Kröhnert, S. Ulbrich, T. Asfour, G. Metta, R. Dillmann, and G. Sandini, "Simox: A robotics toolbox for simulation, motion and grasp planning," in *International Conference on Intelligent Autonomous Systems (IAS)*, 2012, pp. 585–594.

[14] N. Vahrenkamp, T. Asfour, G. Metta, G. Sandini, and R. Dillmann, "Manipulability analysis," in *Proceedings of IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, 2012, pp. 568–573.

[15] D. A. Rosenbaum, R. J. Meulenbroek, J. Vaughan, and C. Jansen, "Posture-based motion planning: applications to grasping." *Psychological review*, vol. 108, no. 4, p. 709, 2001.

[16] M. Katayama and H. Hasuura, "Optimization principle determines human arm postures and" comfort"," in *SICE 2003 Annual Conference (IEEE Cat. No. 03TH8734)*, 2003.

[17] M. Wächter, S. Ottenhaus, M. Kröhnert, N. Vahrenkamp, and T. Asfour, "The ArmarX statechart concept: Graphical programming of robot behaviour," *Frontiers in Robotics and AI*, vol. 3, p. 33, 2016.

[18] N. Vahrenkamp, M. Do, T. Asfour, and R. Dillmann, "Integrated grasp and motion planning," in *IEEE International Conference on Robotics and Automation (ICRA)*, Anchorage (AK), 2010, pp. 2883–2888.

[19] K. Welke, D. Schiebener, T. Asfour, and R. Dillmann, "Gaze selection during manipulation tasks," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2013, pp. 652–659.