Grasp pose estimation in human-robot manipulation tasks using wearable motion sensors

Denis Ćehajić, Sebastian Erhart and Sandra Hirche

***Abstract*— Knowledge of the human grasp pose is crucial in common control schemes for human-robot object manipulation tasks. Biased estimates of the grasp pose cause undesired interaction wrenches on the human partner, which disturbs the interaction and the recognition of motion intention. A use of wearable motion sensors for tracking the human motion facilitates the grasp pose estimation without a global sensing system. This paper presents an approach for estimating an unknown grasp pose of the human using wearable motion sensors while minimizing undesired interaction wrenches applied to the human. A condition necessary for convergence of the estimator together with appropriate robot motion strategies are provided. Estimation of relative orientation and displacement is performed online and based on minimizing the error in the least-square sense. The estimation process does not rely on a global sensing system and it considers only the measurements of the velocity and acceleration of the cooperating partners in their respective local frames. The approach is experimentally evaluated in a physical human-robot interaction scenario.**

I. Introduction

There is a vast interest in scenarios where robots are closely interacting with humans and performing tasks cooperatively. A broad set of challenges arises in tasks where a human and a robot are physically interacting. Safety and robustness issues are commonly in the focus of human-robot interaction (HRI) strategies [1]. Typical applications of physical HRI include human-robot object manipulation (as in Fig. 1) used in manufacturing, construction, and logistics.

In cooperative manipulation tasks a human and a robot are tightly coupled through the object and a coordination strategy is employed to specify the reactive/compliant behaviour of the robot. The corresponding control strategy involves kinematics of the manipulation task: either the human imposes a desired trajectory and the robot needs to continuously adapt during task execution [2] or the human and robot agree in advance on a desired trajectory [3]. Thus, a biased estimate of the grasp pose, i.e. the relative kinematics between the human hand and the robotic end-effector, causes undesired interaction wrenches between partners [4] and results in non-matching robot and human trajectories. The undesired interaction wrenches bias human intention recognition schemes based on interaction wrenches [3]. An incompatible robot trajectory leads to collisions of the robot with its environment, in particular when avoiding obstacles or passing narrow passages. Thus, the robot requires accurate knowledge of the human's grasp pose.

D. Ćehajić, S. Erhart and S. Hirche are with the Chair of Information-oriented Control, Technische Universität München, Munich, Germany. {denis.cehajic, erhart, hirche}@tum.de

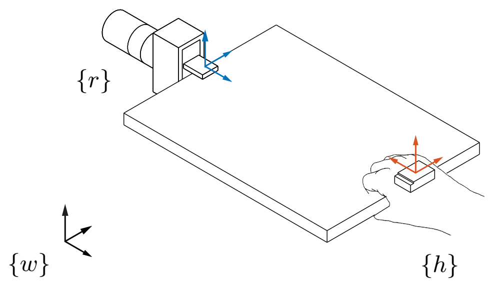

Fig. 1. Human-robot cooperative object manipulation scenario

Conventional approaches in physical HRI assume their relative grasp kinematics is known [1], [3], [5]. This is clearly an oversimplification since the human might not grasp the object at a predefined location or the human might regrasp it during the task. Visually tracking the human hand pose may not have sufficient accuracy due to varying lighting conditions and potential occlusions. In related work in multi-robot cooperative manipulation, the relative kinematics between the robotic end-effectors is identified based on the available end-effector motion signals [6], [7]. However, it is not straightforward to extend the existing approaches to human-robot manipulation. On the one side, the human is unaware of the required trajectory for the identification. On the other hand, the robot solely executing the identification-relevant trajectory may lead to undesired wrenches at the human hand unless human presence is considered.

Wearable motion sensors such as inertial measurement units (IMU's) facilitate grasp pose estimation in HRI. However, so far there are only few works addressing this issue. The estimation of the relative position and orientation of the human is treated by [8]: the human is assumed to behave as a virtual passive revolute joint and the robot motion is driven along the unconstrained circular direction around the human. However, only a planar task is modelled and it is furthermore not clear how to achieve the persistency of excitation (PE) condition [9], assuring the convergence of the grasp pose estimates. Identifying unknown kinematic properties in a physical HRI scenario and deriving suitable robot motion with the human in the loop is still an open issue.

The main contribution of this paper is a strategy for identifying the unknown grasp pose of the human while minimizing the undesired interaction wrenches applied to the human. We model the task and derive an online estimation strategy of the relative kinematics by minimizing the error in the least-square sense: the relative orientation is estimated using the quaternion estimator and the relative displacement using the recursive least squares algorithm. We characterize the persistency of excitation condition involving motion inputs. We formulate a strategy in the case of a static human pose and validate the approach with real-time measurements in a human-robot manipulation setting.

The remainder of the paper is structured as follows. Sec. II formulates the problem. Sec. III presents the estimation strategy and discusses its convergence. Sec. IV treats the robot trajectory selection and Sec. V evaluates the approach.

II. Problem statement

Consider a scenario where a human, denoted with a coordinate frame $\{h\}$, and a robot, $\{r\}$, cooperatively manipulate a rigid object as depicted in Fig. 1. The grasps of the human and the robot are assumed to be rigid, i.e. no slippage occurs during manipulation. The end-effector pose is represented by

$$\boldsymbol{x}_i = \begin{bmatrix} \boldsymbol{p}_i \\ \boldsymbol{q}_i \end{bmatrix} \tag{1}$$

with $\boldsymbol{x}_i \in SE(3)\ \forall i \in h, r$ containing the human/robot translation $\boldsymbol{p}_i \in \mathbb{R}^3$ in the world/inertial frame $\{w\}$, and orientation $\boldsymbol{q}_i \in SO(3)$ represented by a unit quaternion, i.e. $\boldsymbol{q}_i = [\eta_i\ \boldsymbol{\epsilon}_i^T]^T$ with $\eta_i \in \mathbb{R}$ as the real part and $\boldsymbol{\epsilon}_i = [\epsilon_{i1}\ \epsilon_{i2}\ \epsilon_{i3}]^T \in \mathbb{R}^3$ as the imaginary part of the quaternion. A rotation matrix $\boldsymbol{R}(\boldsymbol{q}_i) \in \mathbb{R}^{3\times3}$ of a quaternion $\boldsymbol{q}_i$ is defined as $\boldsymbol{R}(\boldsymbol{q}_i) = (2\eta_i^2 - 1)\boldsymbol{I}_3 + 2\boldsymbol{S}(\boldsymbol{\epsilon}_i) + 2\boldsymbol{\epsilon}_i\boldsymbol{\epsilon}_i^T$ with $\boldsymbol{S} \in \mathbb{R}^{3\times3}$ as a skew-symmetric matrix. Let $\dot{\boldsymbol{x}}_i = [\boldsymbol{v}_i^T\ \boldsymbol{\omega}_i^T]^T \in se(3)$ be a twist vector containing the linear velocity $\boldsymbol{v}_i \in \mathbb{R}^3$, and the angular velocity $\boldsymbol{\omega}_i \in \mathbb{R}^3$ of the human/robot, and let $\ddot{\boldsymbol{x}}_i = [\dot{\boldsymbol{v}}_i^T\ \dot{\boldsymbol{\omega}}_i^T]^T$ contains linear and angular accelerations.

The position and orientation of the human and the robot expressed in the world frame $\{w\}$ are related by

$$\boldsymbol{p}_h = \boldsymbol{p}_r + \boldsymbol{R}(\boldsymbol{q}_r)\,{}^r\boldsymbol{r}_h \tag{2}$$

$$\boldsymbol{q}_h = \boldsymbol{q}_r \otimes {}^r\boldsymbol{q}_h \tag{3}$$

where the relative kinematics is represented by ${}^r\boldsymbol{r}_h \in \mathbb{R}^3$ as the relative displacement of $\{h\}$ with respect to $\{r\}$, and ${}^r\boldsymbol{q}_h$ as the relative orientation of $\{h\}$ with respect to $\{r\}$ (as in Fig. 2). Symbol $\otimes$ denotes the quaternion multiplication.

We employ a human-centered task specification where the desired trajectory of the human hand $\boldsymbol{x}_h^d(t)$ is known as in [3] with constant stiffness (eg. in [10]). In order to continuously adapt the robot motion, a desired trajectory $\boldsymbol{x}_r^d(t)$ of the robot must be computed with respect to that of the human using (2) and (3). The relative kinematics of the human and robot end-effector, i.e. the relative translation and orientation are critical.

The model relating the displacement of the human wrist pose to the applied wrench as in [10] is assumed

$$\boldsymbol{x}_h^d - \boldsymbol{x}_h = \begin{bmatrix} \boldsymbol{K}_{h,t}^{-1} & \boldsymbol{0}_3 \\ \boldsymbol{0}_3 & \boldsymbol{K}_{h,r}^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{f}_h \\ \boldsymbol{t}_h \end{bmatrix} \tag{4}$$

wherein $\boldsymbol{K}_{h,t}$ and $\boldsymbol{K}_{h,r}$ are the translational and rotational stiffness of the human wrist respectively, and $\boldsymbol{u}_h = [\boldsymbol{f}_h^T\ \boldsymbol{t}_h^T]^T$ is the wrench at the human wrist composed of the force $\boldsymbol{f}_h \in \mathbb{R}^3$ and torque vector $\boldsymbol{t}_h \in \mathbb{R}^3$. As a result of (4), the wrench at the human wrist diminishes when the desired and the actual wrist pose coincide $\boldsymbol{x}_h^d = \boldsymbol{x}_h$. Human and robot interaction wrenches are related by

$$\begin{bmatrix} \boldsymbol{f}_h \\ \boldsymbol{t}_h \end{bmatrix} = \begin{bmatrix} \boldsymbol{R}({}^r\boldsymbol{q}_h)^T & \boldsymbol{0}_3 \\ -\boldsymbol{R}({}^r\boldsymbol{q}_h)^T\boldsymbol{S}({}^r\boldsymbol{r}_h) & \boldsymbol{R}({}^r\boldsymbol{q}_h)^T \end{bmatrix} \begin{bmatrix} \boldsymbol{f}_r \\ \boldsymbol{t}_r \end{bmatrix} \tag{5}$$

Given the desired trajectory of the human, $\boldsymbol{x}_h^d(t)$, our aim is to define a desired robot trajectory, $\boldsymbol{x}_r^d(t)$, such that the relative grasp pose estimates, ${}^r\hat{\boldsymbol{r}}_h$ and ${}^r\hat{\boldsymbol{q}}_h$, converge to their true values, ${}^r\boldsymbol{r}_h$ and ${}^r\boldsymbol{q}_h$, i.e.

$${}^r\hat{\boldsymbol{r}}_h \to {}^r\boldsymbol{r}_h \quad \text{and} \quad {}^r\hat{\boldsymbol{q}}_h \to {}^r\boldsymbol{q}_h \quad \text{as } t \to \infty \tag{6}$$

while minimizing the undesired interaction wrenches, i.e.

$$\min_{x_r^d} \|\boldsymbol{u}_h(\boldsymbol{x}_r^d)\|^2 \tag{7}$$

Achieving both (6) and (7) is challenging since the input motions have to satisfy the persistency of excitation condition in order to achieve the estimator convergence and hence, successful identification of the relative kinematics. Performing such motions causes an undesired interaction wrenches when contradicting the human motions as illustrated by (4).

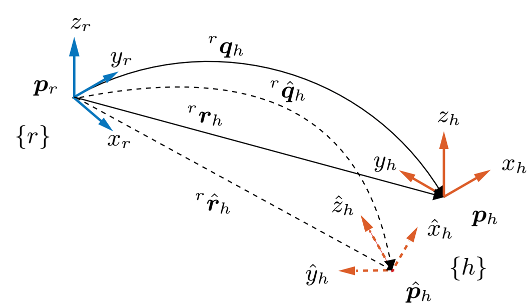

Fig. 2. True ${}^r\boldsymbol{r}_h, {}^r\boldsymbol{q}_h$ and estimated ${}^r\hat{\boldsymbol{r}}_h, {}^r\hat{\boldsymbol{q}}_h$ relative displacement and orientation between the human and robot frames.

A. Manipulation task kinematics

When grasped rigidly, the object mutually constrains the translational and rotational motion of the human $\{h\}$ and the robot $\{r\}$. This implies a constant relative displacement of the human hand with respect to the robot end-effector, i.e.

$${}^r\boldsymbol{r}_h = \text{const.} \tag{8}$$

Differentiating (2) and defining $\boldsymbol{r}_h = \boldsymbol{R}(\boldsymbol{q}_r)\,{}^r\boldsymbol{r}_h$ we obtain expressions for translational velocity

$$\boldsymbol{v}_h = \boldsymbol{v}_r + \boldsymbol{\omega}_r \times \boldsymbol{r}_h \tag{9}$$

and for linear acceleration by differentiating again the previous expression

$$\dot{\boldsymbol{v}}_h = \dot{\boldsymbol{v}}_r + \dot{\boldsymbol{\omega}}_r \times \boldsymbol{r}_h + \boldsymbol{\omega}_r \times (\boldsymbol{\omega}_r \times \boldsymbol{r}_h) \tag{10}$$

The relative orientation between the robot and the human frame is also constrained

$${}^r\boldsymbol{q}_h = \text{const.} \tag{11}$$

By respective differentiation of (3), we obtain the expression for relative angular velocity and acceleration

$$\boldsymbol{\omega}_h = \boldsymbol{\omega}_r \quad \text{and} \quad \dot{\boldsymbol{\omega}}_h = \dot{\boldsymbol{\omega}}_r \tag{12}$$

The linear and angular velocities of the human and the robot in their own local frames are related by

$$ {}^{r}\dot{\boldsymbol{x}}_r = \begin{bmatrix} \boldsymbol{R}({}^{r}\boldsymbol{q}_h) & \boldsymbol{S}({}^{r}\boldsymbol{r}_h)\boldsymbol{R}({}^{r}\boldsymbol{q}_h) \\ \boldsymbol{0}_3 & \boldsymbol{R}({}^{r}\boldsymbol{q}_h) \end{bmatrix} {}^{h}\dot{\boldsymbol{x}}_h \tag{13} $$

The relative linear acceleration is

$$ {}^{r}\dot{\boldsymbol{v}}_r = \boldsymbol{R}({}^{r}\boldsymbol{q}_h)^{h}\dot{\boldsymbol{v}}_h + {}^{r}\dot{\boldsymbol{\omega}}_r \times {}^{r}\boldsymbol{r}_h + {}^{r}\boldsymbol{\omega}_r \times ({}^{r}\boldsymbol{\omega}_r \times {}^{r}\boldsymbol{r}_h) \tag{14} $$

We aim at defining $\boldsymbol{x}_r^d(t)$ given $\boldsymbol{x}_h^d(t)$. Since $\boldsymbol{x}_r^d = \boldsymbol{x}_r^d({}^{r}\boldsymbol{r}_h, {}^{r}\boldsymbol{q}_h)$ as seen in (2) and (3), the estimator of the relative kinematics is needed. For simplicity, let us denote $\boldsymbol{r} = {}^{r}\boldsymbol{r}_h$ and $\boldsymbol{q} = {}^{r}\boldsymbol{q}_h$.

III. Grasp pose estimation

The parameter identification scheme aims at identifying the unknown quantities in (6), contained in the kinematic constraints. The objective of this section is to find a suitable combination of a sensing system and an estimation strategy.

A. Selecting an appropriate sensing system

Any of the constraint representations (pose, velocity, acceleration) is suitable to conclude on the actual parameters. However, the available sensing equipment on the human and the robot might privilege or limit the use of a certain representation. If the human hand and the robotic end-effector could be tracked by the robot vision or an optical tracking system (yielding $\boldsymbol{x}_h(t)$ and $\boldsymbol{x}_r(t)$ in $\{w\}$), the pose representation appears convenient. However, possible occlusions during grasping and inaccuracies limit the use of human grasp pose tracking by vision. Inertial sensors are more appropriate for tracking the human motion. This requires formulating an estimation strategy at the velocity/acceleration level.

Inertial sensors typically provide measurements of linear acceleration $\dot{\boldsymbol{v}}$ and angular velocity $\boldsymbol{\omega}$ of the motion. In order to formulate an estimator at the velocity level, $\dot{\boldsymbol{v}}$ needs to be integrated in order to obtain $\boldsymbol{v}$, since only $\boldsymbol{\omega}$ is directly available. However, errors occur during integration. We formulate an estimator based on relative angular velocity (13) and linear acceleration (14) expressions and thus avoid the noisy integration. Measurements of the human end-effector (coinciding with the human grasp location) are obtained by inertial sensors placed at the human hand. Robot end-effector measurements are obtained through the forward kinematics.

B. Estimating relative kinematics

Given measurements of the angular velocity and linear acceleration of the human and robot end-effectors, ${}^{h}\boldsymbol{\omega}_h$, ${}^{r}\boldsymbol{\omega}_r$, ${}^{h}\dot{\boldsymbol{v}}_h$, and ${}^{r}\dot{\boldsymbol{v}}_r$, we derive an *online* estimator of the relative kinematics. Let us define error residuals of angular velocities and linear accelerations from eqs. (13) and (14) as

$$ \boldsymbol{e}_{\omega,i} = {}^{r}\boldsymbol{\omega}_{r,i} - \boldsymbol{R}(\hat{\boldsymbol{q}})^{h}\boldsymbol{\omega}_{h,i} \tag{15} $$

$$ \begin{aligned} \boldsymbol{e}_{\dot{v},i} = {}^{r}\dot{\boldsymbol{v}}_{r,i} - \boldsymbol{R}(\hat{\boldsymbol{q}})^{h}\dot{\boldsymbol{v}}_{h,i} - {}^{r}\dot{\boldsymbol{\omega}}_{r,i} \times \hat{\boldsymbol{r}} \\ - {}^{r}\boldsymbol{\omega}_{r,i} \times ({}^{r}\boldsymbol{\omega}_{r,i} \times \hat{\boldsymbol{r}}) \end{aligned} \tag{16} $$

where subscript i denotes estimates/measurements at time t_i.

An optimization problem solving for unknown $\boldsymbol{r}$ and $\boldsymbol{q}$ can be formulated in terms of least squares as in [7]. Solving the multi-variable optimization problem for both kinematic parameters might be difficult, computationally intensive, and hence not applicable online. We decouple the problem of orientation and displacement estimation in the following sense: the orientation estimator minimizing (15) requires no displacement estimate and can be computed separately, the displacement estimator minimizing (16) is computed using estimated orientation. For notational convenience, let us remove superscripts of the local frames, e.g. ${}^{h}\dot{\boldsymbol{\omega}}_h = \dot{\boldsymbol{\omega}}_h$.

Estimating relative orientation: Let us define a cost function using the relative orientation error (15) up to time t_k

$$ J_{\omega,k} = \frac{1}{2}\sum_{i=1}^{k} a_i \left\|\boldsymbol{e}_{\omega,i}\right\|^2 \quad \text{s.t. } \|\hat{\boldsymbol{q}}\| = 1 \tag{17} $$

subject to the quaternion norm unity constraint. Measurements are weighted by $a_i = e^{-\mu_q(t_k - t_i)}\ \forall i = 1, 2, \ldots, k$ with the forgetting factor $\mu_q < 1$, and k as the total number of measurements.

The quaternion estimator solving the Wahba's problem [11] is employed to find the optimal estimate of the relative orientation $\hat{\boldsymbol{q}}^*$ recursively for each sampling time t_s with respect to the cost function (15)

$$ \Phi_{q,k} = \delta_q \Phi_{q,k-1} + \boldsymbol{\phi}_{q,k} \qquad \forall i = 1 \ldots k \tag{18} $$

where $\Phi_{q,k} = \sum_{i=1}^{k} a_i \boldsymbol{\phi}_{q,k}$ and $\boldsymbol{\phi}_{q,k}(\boldsymbol{\omega}_{r,k}\ , \boldsymbol{\omega}_{h,k}) \in \mathbb{R}^{4\times4}$ is a symmetric traceless matrix

$$ \boldsymbol{\phi}_{q,k} = \begin{bmatrix} \boldsymbol{\omega}_{h,k}\boldsymbol{\omega}_{r,k}^T + \boldsymbol{\omega}_{r,k}\boldsymbol{\omega}_{h,k}^T - \boldsymbol{\omega}_{r,k}^T\boldsymbol{\omega}_{h,k}\boldsymbol{I}_3 & \boldsymbol{\omega}_{h,k} \times \boldsymbol{\omega}_{r,k} \\ (\boldsymbol{\omega}_{h,k} \times \boldsymbol{\omega}_{r,k})^T & \boldsymbol{\omega}_{r,k}^T\boldsymbol{\omega}_{h,k} \end{bmatrix} $$

with an initial orientation $\boldsymbol{q}_0$ such that $\|\boldsymbol{q}_0\| = 1$. The weight scalar is $\delta_q = e^{-\mu_q t_s}$, and the factor μ_q is as in (17). The optimal estimate $\hat{\boldsymbol{q}}^*$ is the eigenvector of the largest eigenvalue of $\boldsymbol{\phi}_{q,k}$.

Estimating relative displacement: Relative displacement $\boldsymbol{r}$ is estimated using a standard recursive least-squares estimator minimizing the cost function from (16)

$$ J_{\dot{v},k} = \frac{1}{2}\sum_{i=1}^{k} a_i ||\boldsymbol{e}_{\dot{v},i}||^2 \tag{19} $$

Note that (14) can be expressed as a linear model

$$ \dot{\boldsymbol{v}}_{r,k} - \boldsymbol{R}(\boldsymbol{q})\dot{\boldsymbol{v}}_{h,k} = \boldsymbol{\phi}_{r,k}\boldsymbol{r} \tag{20} $$

with $\boldsymbol{\phi}_{r,k}(\dot{\boldsymbol{\omega}}_{r,k}\ , \boldsymbol{\omega}_{r,k}) \in \mathbb{R}^{3\times3}$ being the regressor matrix

$$ \boldsymbol{\phi}_{r,k} = \boldsymbol{S}(\dot{\boldsymbol{\omega}}_{r,k}) + \boldsymbol{S}(\boldsymbol{\omega}_{r,k})\boldsymbol{S}(\boldsymbol{\omega}_{r,k}) \tag{21} $$

The update of the displacement estimator is

$$ \hat{\boldsymbol{r}}_{k+1} = \hat{\boldsymbol{r}}_k + \boldsymbol{K}_k \boldsymbol{e}_{\dot{v},k} \tag{22} $$

with the error $\boldsymbol{e}_{\dot{v},k}$ as in (16) considering only measurements and the orientation estimate at k, the gain $\boldsymbol{K}_k \in \mathbb{R}^{3\times3}$ and covariance matrix $\boldsymbol{P}_{k+1} \in \mathbb{R}^{3\times3}$ defined as

$$ \begin{aligned} \boldsymbol{K}_k &= \boldsymbol{P}_k\ \boldsymbol{\phi}_{r,k}(\delta_r \boldsymbol{I}_3 - \boldsymbol{\phi}_{r,k}\boldsymbol{P}_k\boldsymbol{\phi}_{r,k})^{-1} \\ \boldsymbol{P}_{k+1} &= \frac{1}{\delta_r}(\boldsymbol{I}_3 + \boldsymbol{K}_k\boldsymbol{\phi}_{r,k})\boldsymbol{P}_k \end{aligned} $$

with the initial displacement $\boldsymbol{r}_0$ and covariance matrix $\boldsymbol{P}_0$, factor δ_r is the weight as in (18) with the forgetting factor μ_r.

C. Persistency of excitation condition

In order for the estimators (18) and (22) to converge minimizing (15) and (16), the input motions need to satisfy the persistency of excitation condition [9]. From a system theoretic point of view this can be interpreted as an identifiability condition of the parameter vector which depends on a specific input for nonlinear systems. In our case the PE condition describes sufficiently "rich" motions for identifying the human grasp pose. The relative displacement estimate $\hat{\boldsymbol{r}}$ will converge towards $\boldsymbol{r}$ if there are positive constants T and α such that

$$\int_t^{t+T} \boldsymbol{\phi}_r^T \boldsymbol{\phi}_r \geq \alpha \boldsymbol{I}_3 \tag{23}$$

is satisfied $\forall t$, with $\boldsymbol{\phi}_r$ being the regressor matrix as in (21). The regressor matrix is composed of the sum and the product of skew-symmetric matrices $\boldsymbol{S}(\boldsymbol{\omega}_r)$ and $\boldsymbol{S}(\dot{\boldsymbol{\omega}}_r)$. Its rank is $\text{rank}(\boldsymbol{\phi}_r) \geq 2$. The corresponding skew-symmetric matrix of a non-zero input motion vector $\boldsymbol{\omega}_r$ has $\text{rank}(\boldsymbol{S}(\boldsymbol{\omega}_r)) = 2$. The associated null-space $\text{null}(\boldsymbol{S}(\boldsymbol{\omega}_r)) = \text{span}(\boldsymbol{\omega}_r)$. The range of $\boldsymbol{S}(\boldsymbol{\omega}_r)$ is the plane orthogonal to $\boldsymbol{\omega}_r$. The same is valid for a non-zero input motion vector $\dot{\boldsymbol{\omega}}_r$. The null-space of the matrix product $\boldsymbol{S}(\boldsymbol{\omega}_r)\boldsymbol{S}(\boldsymbol{\omega}_r)$ in (21) is also $\text{span}(\boldsymbol{\omega}_r)$ with an analogue result for the range.

This means that in the case when $\boldsymbol{\omega}_r$ and $\dot{\boldsymbol{\omega}}_r$ are non-collinear, $\boldsymbol{\phi}_r$ has full rank, i.e. $\text{rank}(\boldsymbol{\phi}_r) = 3$, since the range of the sum of $\boldsymbol{S}(\boldsymbol{\omega}_r)$ and $\boldsymbol{S}(\dot{\boldsymbol{\omega}}_r)$ span $\mathbb{R}^3$. Thus $\boldsymbol{\phi}_r^T \boldsymbol{\phi}_r$ is full rank satisfying (23), i.e. the motion of the system is persistently exciting (22). The motions also persistently excite the orientation estimator (18).

Fulfilling the PE condition having a human in the loop is challenging. Next section discusses how to achieve/guarantee such motions given $\boldsymbol{x}_h^d(t)$.

IV. Trajectory selection

In this section we are interested in finding a suitable robot trajectory aiming at identifying the grasp pose and minimizing the undesired interaction wrenches. We analyse the case when the human motion is static. Although not studied in depth, we illustrate strategies and give intuitive explanations in cases of dynamic human behaviour.

A. Static human pose

Let us firstly consider the case when the desired human wrist trajectory is static, i.e. the human wrist is at rest. In this case the human hand pose is $\boldsymbol{x}_h^d(t) = \text{const.}$ and any force/torque applied by the robot in view of the grasp pose identification is considered as a perturbation to the human.

In order to guarantee the parameter convergence (cf. (23)), the robot needs to induce a motion of the object with a non-collinear and a non-zero angular velocity, i.e. $\boldsymbol{\omega}_r \neq 0$. Since $\boldsymbol{\omega}_r = \boldsymbol{\omega}_h$, it follows that the torque at the human wrist $\boldsymbol{t}_h$ cannot be avoided. The torque produced by the robot to induce an angular velocity is determined by the desired amplitude of the motion that the robot performs.

Inducing an angular motion by the robot requires considering the most appropriate axis of rotation. Since our aim is not to perturb, i.e. move the human wrist satisfying $\boldsymbol{v}_h = 0$ (and hence minimal force disturbance $\boldsymbol{f}_h$ in (7)), let us consider the expression of the kinematic constraint (9)

$$\boldsymbol{v}_h = \underbrace{\begin{bmatrix} \boldsymbol{I}_3 & \boldsymbol{S}(\boldsymbol{r}_h) \end{bmatrix}}_{\boldsymbol{M}} \begin{bmatrix} \boldsymbol{v}_r \\ \boldsymbol{\omega}_r \end{bmatrix} \tag{24}$$

By choosing an appropriate $\dot{\boldsymbol{x}}_r \in \text{null}(\boldsymbol{M})$, the translational velocity of the human is $\boldsymbol{v}_h \equiv 0$. The null-space of the matrix $\boldsymbol{M}$ such that $\boldsymbol{M}\dot{\boldsymbol{x}}_r = 0$ is

$$\text{null}(\boldsymbol{M}) = \text{span}\left(\begin{bmatrix} \boldsymbol{S}(\boldsymbol{r}_h)^T \\ \boldsymbol{I}_3 \end{bmatrix} \right) \tag{25}$$

By choosing the human wrist ${}^r\boldsymbol{r}_h$ as the axis of rotation there is no change of human wrist position, i.e. $\boldsymbol{p}_h = \int \boldsymbol{v}_h = \text{const.}$ and $\boldsymbol{f}_h = 0$ since $\boldsymbol{p}_h^d - \boldsymbol{p}_h = 0$. The robot desired trajectory at the velocity level is defined as

$$\dot{\boldsymbol{x}}_r^d(t) = \begin{bmatrix} \boldsymbol{v}_r^d(t) \\ \boldsymbol{\omega}_r^d(t) \end{bmatrix} = \begin{bmatrix} \hat{\boldsymbol{r}} \times \boldsymbol{\omega}_r^d(t) \\ \boldsymbol{\omega}_r^d(t) \end{bmatrix} \tag{26}$$

where $\boldsymbol{\omega}_r^d(t)$ is defined such that at two consecutive time instances, the vectors are non-collinear satisfying (23).

B. Planar human motion

Consider now the case when the human is performing a task-relevant translational movement in a plane in $SE(3)$. Although more challenging due to the human dynamic behaviour, the planar restriction enables identification of a suitable identification strategy. Desired robot motion needs to assure both tracking the human translational velocity and inducing an identification-relevant motion.

Since we have 3D velocity/acceleration measurements of the human and robot end-effectors, a projected plane spanned by the human motions $\boldsymbol{x}_h^d(t)$ can be defined with the normal vector resulting from the cross-product of two non-collinear measurements. In order to achieve (7), the human translational velocity vector can be chosen as the axis of rotation for inducing angular motions satisfying (23). Since angular motions are induced around a single motion direction, only the subspace of the grasp parameters are identified, i.e. two-dimensional components of the 3D grasp pose. In order to estimate the complete parameter space, a rotational motion around two non-collinear (piecewise constant) translational velocity vectors of the human wrist $\boldsymbol{v}_h$ is needed.

C. 6D human motion

Consider now human motions spanning the complete 6D motion space. In this case, the human either performs angular motions that are already persistently exciting the estimator such that (23) is satisfied. Alternatively, the task is approximated by the strategies used in the cases of the static motion or by identifying the planes of the human motion. This challenging scenario requires an in-depth analysis of potential subspaces of the human motion and we therefore focus on the case when the human motion is static.

V. Experimental evaluation

A. Experimental setup

The proposed approach is evaluated in a human-robot manipulation experiment. The robot and the human rigidly grasp the object with a human wrist at rest (Fig. 3). A 7 DoF robotic manipulator is equipped with a two-fingered Schunk PG70 gripper at the end-effector. An Xsens IMU sensor is placed at the human wrist for measuring the motions. The interaction wrenches are measured by a 6 DoF JR3 force/torque sensor. The object is a rigid rectangular aluminium frame (1.0×0.8 m) weighted 2 kg.

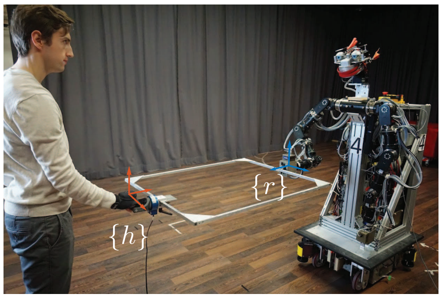

Fig. 3. Experimental setup and frame alignment: the x-axis of $\{h\}$ and $\{r\}$ point towards each other, the z-axis of both frames point upwards.

The robot control scheme consists of an impedance-based controller (enabling the robot compliant behaviour) and a joint-space position controller with inverse kinematics (enabling the robot reference trajectory tracking in task-space). This control architecture is commonly used in physical HRI to assure human safety. Parameters of the impedance controller are set to $\boldsymbol{M} = \left[10\boldsymbol{I}_3 \text{ kg},\ \boldsymbol{0}_3; \boldsymbol{0}_3,\ 0.5\boldsymbol{I}_3 \text{ kgm}^2\right]$ for inertia, $\boldsymbol{D} = \left[180\boldsymbol{I}_3 \text{ Ns/m},\ \boldsymbol{0}_3; \boldsymbol{0}_3,\ 10\boldsymbol{I}_3 \text{ Nms/rad}\right]$ for damping, and $\boldsymbol{K} = \left[300\boldsymbol{I}_3 \text{ N/m},\ \boldsymbol{0}_3; \boldsymbol{0}_3,\ 50\boldsymbol{I}_3 \text{ N/m}\right]$ for stiffness. The sampling rate of the estimation and control is set to $t_s = 1$ ms. A common reference frame is not required. True values of the human grasp pose are $\boldsymbol{r} = [1.11\ 0.16\ 0.08]^T$ (m) and $\boldsymbol{q} = [0\ 0\ 0\ 1]^T$.

B. Results and discussion

In the first run, a robot reference trajectory is chosen as $\dot{\boldsymbol{x}}_r^{\text{ref}} = [\boldsymbol{v}_r^{\text{ref}}\ \boldsymbol{\omega}_r^{\text{ref}}]^T$ with $\boldsymbol{\omega}_r^{\text{ref}} = [10^\circ\ 40^\circ\ 20^\circ]^T \frac{\pi \text{ rad}}{180^\circ} \sin(t)$, to satisfy (23), and $\boldsymbol{v}_r^{\text{ref}} = \boldsymbol{r}_0 \times \boldsymbol{\omega}_r^{\text{ref}}$, with the initial $\boldsymbol{r}_0 = [0.5\ 0\ 0]^T$ and $\boldsymbol{q}_0 = [0.5\ 0.5\ 0.5\ 0.5]^T$. The estimator factors are set to $\mu_r = 0.9$ and $\mu_q = 0.5$ to account for the noisy measurements of the IMU and robot sensors. The initial covariance is set to $\boldsymbol{P}_0 = \boldsymbol{I}_3$.

The absolute errors of the estimated relative orientation and displacement are shown in Fig. 4. Within ~ 10 s, the orientation estimator approaches the true quaternion value and after ~ 30 s. it reaches the steady-state. The estimated quaternion after 1 min. is $\hat{\boldsymbol{q}} = [0.012\ 0.001\ 0.026\ 0.999]^T$. The displacement estimation is influenced by the estimated orientation $\hat{\boldsymbol{q}}_t$ at time t as reflected through (20). The initial error of the estimator is significantly reduced after 1 min. The estimator reaches the estimate $\hat{\boldsymbol{r}} = [1.03\ 0.20\ 0.10]^T$. The estimators show similar behaviour with different initial guesses. The estimator's convergence speed is influenced by the choice of the selected reference trajectory which is subject to the robot velocity limits. Although the simulation results show fast convergence with the simulated input, the mechanical limitations prevents us from experimenting with higher end-effector velocities. Inertial sensor noise used for human motion tracking also affects the performance of the estimators: noise is $0.05^\circ/s\sqrt{Hz}$ for $\boldsymbol{\omega}_h$ and $0.002^\circ/s^2\sqrt{Hz}$ for $\dot{\boldsymbol{v}}_h$. The noise occurring in the orientation calculated from the gyroscope additionally affects the measurements of $\dot{\boldsymbol{v}}_h$ in which the gravity is firstly compensated prior to displacement estimation by $\dot{\boldsymbol{v}}_h = \dot{\boldsymbol{v}}_h^{\text{raw}} - \boldsymbol{R}(\boldsymbol{q}_h^{\text{raw}})\ [0\ \ 0\ \ 9.81]^T$ with $\dot{\boldsymbol{v}}_h^{\text{raw}}$ and $\boldsymbol{q}_h^{\text{raw}}$ being the raw measurements of the inertial sensor. An angle error of 0.1° in orientation causes 1.71 m/s^2 error in $\dot{\boldsymbol{v}}_h$. Since the orientation estimation depends only on measured $\boldsymbol{\omega}_h$, its performance is superior with respect to the displacement estimator performance which depends on both the measured $\boldsymbol{\omega}_h$ and $\dot{\boldsymbol{v}}_h$. Similar results of the displacement estimation are obtained by performing the identification with the known relative orientation. Strategies for improving the estimators are currently being investigated.

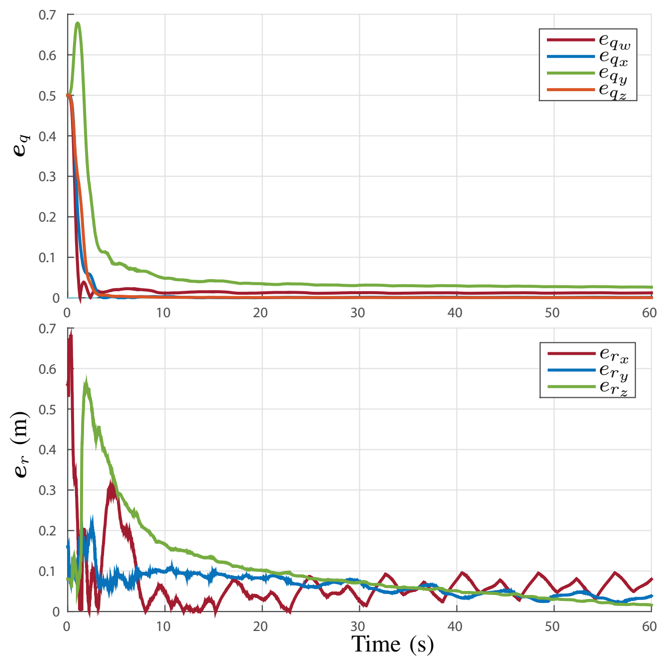

Fig. 4. Absolute error of the relative human grasp estimation: (top) orientation error $\boldsymbol{e}_q = |\boldsymbol{q} - \hat{\boldsymbol{q}}|$ with individual errors $e_{q_i} = |q_i - \hat{q}_i|$, (bottom) translation error $\boldsymbol{e}_r = |\boldsymbol{r} - \hat{\boldsymbol{r}}|$ with $e_{r_i} = |r_i - \hat{r}_i|$.

In the second and third run we show the effect of the biased and the estimated grasp pose seen through the interaction force applied on the human wrist. It is minimized by choosing the human grasp pose as the axis of rotation. As pointed out, the interaction torque is determined by the amplitude of the applied rotational motion by the robot. We focus on showing the effects of biased human grasp displacement on the human interaction force. We set two different values of the relative grasp position. The resulting interaction force is shown in Fig 5. In the case of the initial/biased guess $\boldsymbol{r}_0 = \boldsymbol{r}_{\text{bias}} = [0.5\ 0\ 0]^T$ which could be an incorrectly measured pose, there exists a disturbance occurring during the oscillatory motion of the robot. Measurements show the force fluctuation ranges approximately by 1-3 N along x-axis, from -3.5 to 1 N along y-axis, and 1.1 N along z-axis. However, in the case of $\hat{\boldsymbol{r}} = [1.03\ 0.20\ 0.10]^T$, since the estimate $\hat{\boldsymbol{r}}$ is closer to its true value, the force applied on the human is smaller than in the case of the biased estimate. The force in x-axis is the largest due to the largest estimation error compared to y and z-axis. The disturbance on the human is fluctuating approximately from -0.6 to 1.4 N in x, from -0.2 to 0.5 N in y, and from -0.3 N to 0.07 N in z-axis.

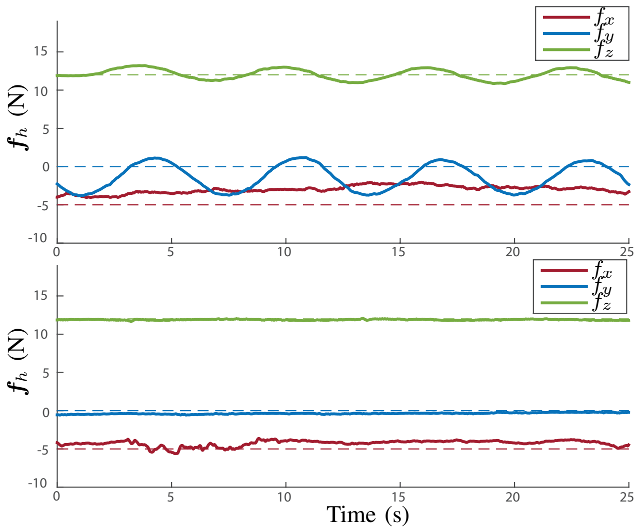

Fig. 5. Interaction force on the human hand: (top) with $\boldsymbol{r}_{\text{bias}} = [0.5\ 0\ 0]^T$ (m), (bottom) with $\hat{\boldsymbol{r}} = [1.03\ 0.20\ 0.10]^T$ (m). Dotted lines denote forces exerted by the human partner when no motion is incurred. The z-axis shows the force that the human applies to hold the object against the gravity and the small negative force in the x-axis represents pulling force from the robot.

In this paper the case of the static human pose is treated with no slippage and regrasp phases. The results are applicable to a wider range of problems where a human and a robot are physically coupled and unknown kinematic parameters exist. During our experiments the human was instructed to be passive at wrist in which the stiffness coefficients were expected to be stable. In a more veridical human impedance model, however, the joint stiffness of the human limbs are known to be time-varying, particularly when the human is actively involved in the task [12]. By assuming a desired human trajectory with a known path (or a pattern) and velocity the proposed approach can be applied on more complex human motions. This results in further assumptions on the task. Furthermore, slippage of the grip is commonly encountered in dexterous tasks despite that we assumed the rigid grasp in our study. The slippage effects can be treated by modelling the human grasp with more complex contact models or using a probabilistic approach where changes in the human grasp can be expressed in terms of uncertainty. Moreover, by improving the estimation convergence time the phases of regrasp can be detected. Lastly, here the reference trajectory is chosen purely for identifying the unknown grasp pose. In a real-world scenario a task-relevant trajectory can be adopted such as moving an object (with unknown human grasp pose) from point A to point B. Techniques of performing both the identification and the manipulation task simultaneously are an ongoing development.

VI. Conclusion

This paper investigated an approach for estimating the human grasp pose using wearable motion sensors in human-robot manipulation tasks. The proposed estimation strategy identifies the unknown relative displacement and orientation of the human grasp. We treat the persistency of excitation condition for the estimator convergence. We show that in the case of a static human pose the undesired interaction force on the human is minimized when the robot motion is performed around the humans wrist. Experiments validating the proposed approach are performed. Our future work will consider more complex human motions and simultaneous identification and task-relevant manipulation.

Acknowledgement

The authors would like to thank M. Lang for his work within his thesis and S. Endo for his helpful insights. The research leading to these results has received funding from the European Union Seventh Framework Programme FP7/2007-2013 under grant agreement n° 601165 of the project "WEARHAP - WEARable HAPtics for humans and robots".

References

[1] A. Bussy, A. Kheddar, A. Crosnier, and F. Keith, "Human-humanoid haptic joint object transportation case study," in *Intelligent Robots and Systems, IEEE/RSJ International Conference on*, 2012, pp. 3633–3638.

[2] E. Gribovskaya, A. Kheddar, and A. Billard, "Motion learning and adaptive impedance for robot control during physical interaction with humans," in *Robotics and Automation, 2011 IEEE International Conference on*, 2011, pp. 4326–4332.

[3] A. Moertl, M. Lawitzky, A. Kucukyilmaz, M. Sezgin, C. Basdogan, and S. Hirche, "The role of roles: Physical cooperation between humans and robots," *International Journal of Robotics Research*, vol. 31, no. 13, pp. 1657–1675, 2012.

[4] S. Erhart, D. Sieber, and S. Hirche, "An impedance-based control architecture for multi-robot cooperative dual-arm mobile manipulation," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2013, pp. 315–322.

[5] K. Kosuge, H. Kakuya, and Y. Hirata, "Control algorithm of dual arms mobile robot for cooperative works with human," in *Systems, Man, and Cybernetics, IEEE International Conference on*, vol. 5, 2001, pp. 3223–3228.

[6] S. Erhart and S. Hirche, "Adaptive force/velocity control for multi-robot cooperative manipulation under uncertain kinematic parameters," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2013, pp. 307–314.

[7] F. Aghili, "Adaptive control of manipulators forming closed kinematic chain with inaccurate kinematic model," *Mechatronics, IEEE/ASME Transactions on*, vol. 18, no. 5, pp. 1544–1554, 2013.

[8] Y. Karayiannidis, C. Smith, F. Vina, and D. Kragic, "Online kinematics estimation for active human-robot manipulation of jointly held objects," in *Intelligent Robots and Systems, IEEE/RSJ International Conference on*, 2013, pp. 4872–4878.

[9] L. Ljung, *Characterization of the concept of 'persistently exciting' in the frequency domain.* Lund Inst. of Technology, Division of Automatic Control, 1971.

[10] T. Tsumugiwa, R. Yokogawa, and K. Hara, "Variable impedance control based on estimation of human arm stiffness for human-robot cooperative calligraphic task," in *Robotics and Automation, IEEE International Conference on*, vol. 1, 2002, pp. 644–650.

[11] M. D. Shuster and S. Oh, "Three-axis attitude determination from vector observations," *Journal of Guidance, Control, and Dynamics*, vol. 4, no. 1, pp. 70–77, 1981.

[12] E. Burdet, G. Ganesh, C. Yang, and A. Albu-Schäffer, "Interaction force, impedance and trajectory adaptation: by humans, for robots," in *Experimental Robotics*. Springer, 2014, pp. 331–345.