

Detection of Collaboration and Collision Events during Contact Task Execution

Felix Franzel², Thomas Eiband^{1,2}, and Dongheui Lee^{1,2}

Abstract—This work introduces a contact event pipeline to distinguish task-contact from Human-Robot interaction and collision during task execution. The increasing need for close proximity physical human-robot interaction (pHRI) in the private, health and industrial sector demands for new safety solutions. One of the most important issues regarding safe collaboration is the robust recognition and classification of contacts between human and robot. A solution is designed, that enables simple task teaching and accurate contact monitoring during task execution. Besides an external force and torque sensor, only proprioceptive data is used for the contact evaluation. An approach based on demonstrated task knowledge and the offset resulting from human interaction is designed to distinguish contact events from normal execution by a contact event detector. A contact type classifier implemented as Support Vector Machine is trained with the identified events. The system is set up to quickly identify contact incidents and enable appropriate robot reactions. An offline evaluation is conducted with data recorded from intended and unintended contacts as well as examples of task-contacts like object manipulation and environmental interactions. The system’s performance and its high responsiveness are evaluated in different experiments including a real world task.

I. INTRODUCTION

Collaborative robots or Cobots are a minor but steadily increasing sector in robotics. The International Federation of Robotics (IFR) predicts a compound annual growth rate between 50% and 60% [1]. With the typical industrial robot working in evacuated workspaces, there is a steadily increasing demand for more and more tasks to be performed with the joined effort of human operator and robot. This collaboration in industrial surroundings can be of different forms. The IFR specifies different levels of collaborative applications [2], ranging from simple coexistence of human and robot in the workspace to responsive collaboration between them. In combining the best abilities from both partners, collaborative robotics can automate tasks that could not be automated so far and thereby enhance productivity.

Safety of the worker, such as summarized by the ISO 10218-1:2011 standard, is thereby the most pressing objective. Different perspectives of safe human robot interaction are analyzed in numerous prior works. Where some analyze the possible injuries or inflicted pain [3]–[5], others propose safety strategies [6], actuation mechanisms [7], interaction

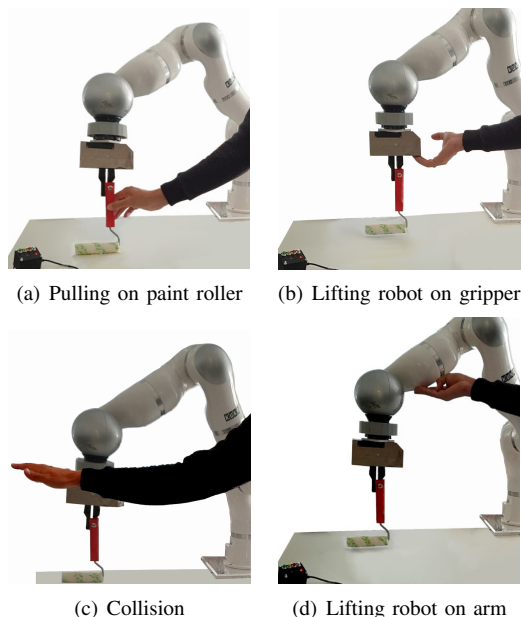


Fig. 1. Four examples of interactions during the experiment to test the contact type prediction performance of the contact event pipeline.

control schemes [8], collision detection [9], reaction strategies to collisions [10] or even complete contact avoidance during interaction [11]. In general, when humans want to work with robots in close proximity, especially when direct interaction contacts are possible, complete safety cannot be guaranteed without sacrificing productivity. A monitoring system must therefore classify such human-robot contact situations to avoid severe injuries. These monitoring systems can be based on camera, motion or tactile sensors [9]–[13], or a combination of them.

For Sequential Collaboration like in [14], human interaction is often confined to the workpiece. Also enabling interaction with the robot arm itself makes it feel more natural to the operator. It is thus desirable to enable all intuitive types of Physical Human Robot Interaction (pHRI) such as guidance of the robot when a task is not executed correctly, reshaping of the trajectory to avoid obstacles in the path, and hindering of the robot’s movement when the operator is in the way [15]. Besides these intended human interactions, contacts can occur in unintended ways if the operator is not aware of the robot’s designated movement or if the robot behaves other than expected. These collisions of many possible forms must be quickly identified to guarantee the worker’s safety. A collaborative human-robot system must thus be able to identify different possible human interactions

¹ German Aerospace Center (DLR), Institute of Robotics and Mechatronics, Wessling, Germany

² Technical University of Munich (TUM), Chair of Human-centered Assistive Robotics, Munich, Germany
This work was supported by the Helmholtz Association.
felix.franzel@tum.com, thomas.eiband@dlr.de, dongheui.lee@dlr.de

Collision	Control Contact	Task-Contact
<ul style="list-style-type: none"> ▪ Non-voluntary (human-robot) contact ▪ Unintended (robot-environment) contact ▪ Dynamic interaction or clamping 	<ul style="list-style-type: none"> ▪ Give commands to the robot ▪ Correction of the robot ▪ Inhibiting robot (to avoid collision) ▪ No command inputs 	<ul style="list-style-type: none"> ▪ Related to robot task ▪ Initiated by the robot ▪ Human interaction during task-contact possible

Fig. 2. Contact types as defined by [16].

(examples in Fig. 1) and classify their type. Additionally, the robot must be able to differentiate between contacts that are caused by the task and those that are occurring from human interaction.

Burghart et al. [16] categorize human-robot and environmental contacts in three main classes (Fig. 2), which are *collision*, *control contact* and *task-contact*. An example of control contact is the correction of the end-effector in a placement task or inhibiting the robot to avoid a collision. Task-contact occurs when the specified task requires object handling or when contact between a tool and the environment is desired. In the following, we will stick to this categorization and terminology.

In summary, a monitoring system must always be aware of the state of contact to discriminate between contacts due to the task and contacts due to engagement with the human body. Examples are cameras, motion capturing systems or artificial robot skin. However, these systems are not without faults since problems like occlusion are likely with camera based systems. According to [17] and [1], high technological flexibility is a key economic factor and task-specific sensor tuning like in [14] is inefficient and time-consuming.

In this work, we propose a system configured as a contact event pipeline that can distinguish task-contact from human interaction and collision during Human-Robot Cooperation. The operator is enabled to physically interact with any part of the robot or tool. Furthermore, the user is protected from unintended, harmful contacts that can possibly occur from sharing the same workspace with the robot. Advantageously, these human interactions are recognized even while a task-contact is active at the same time such that an appropriate reaction can be triggered.

The main contributions of this work are: (i) robust detection of human related contacts with the robot based only on one recording of the task as reference sample; (ii) discrimination of task contact and human interaction during free motion and while in contact; (iii) a method for fast contact classification based on newly designed physical features that outperforms the state of the art (reaction time < 0.2 s), which is evaluated both in offline evaluation and in a real world experiment with a KUKA LWR4+ robot and human subjects.

II. RELATED WORK

It can be seen by the vast amount of research in this field, that estimating the contact state of the robot is a critical issue not only for safety reasons but also for close cooperation and task completion. To encounter this issue, contact between the human and the robot might be totally avoided in order to obviate collisions at any time, as presented in [11]. An artificial repulsive force field around the human operator

changes the planned trajectory to push the robot away from the human body at any time. However, a complete avoidance of physical contact is not desirable for our framework, since control contacts could not be enabled.

Therefore, a framework is required that reacts to human contacts accordingly while executing a task. Golz et al. [12] propose an approach that distinguishes between intentional and unintentional physical human-robot contacts. The classification is achieved with independent non-linear support vector machines per robot joint. In comparison, we compute features from a number of joint and Cartesian state space variables to train a common classification model that resembles an overview of the whole robot arm.

Most existing approaches for collision detection rely solely on proprioceptive information and use a moment observer with an experimentally evaluated threshold [18]. However, they do not consider task and control contact explicitly. In [13], contact situations are analyzed with the help of the Spectral Norm Derivative (SND) of force sensor data measured at the end effector or the proprioceptive joint torques. Advantageously, this approach does not rely on any specific model with a prior training phase or data recordings of example interactions. The distinction threshold between intended and unintended human-robot interactions however must be manually defined and might not generalize over different scenarios. Furthermore, no task-contact can be evaluated by this method. Nevertheless, the SND value gives great insights on the frequency changes of a contact event, an important factor according to [19], and will therefore extend our set of classification features used in our data driven approach.

All of the above frameworks do not facilitate the distinction of task-contact and human interaction including the classification of the intent of the contact. In [20], only task contact is detected in the teaching phase without considering other interaction types. A recent work considers task contact and different human interaction types [21] by employing a data-driven approach for contact detection and evaluation, which is constructed as a three-stage pipeline based on two Recurrent Neural Networks. An estimate of the wrench exerted by the human on a contact point along the robot structure is used to first identify contact events and sequentially classify the contact type. The reference data is retrieved from Gaussian Mixture Models trained by previously recorded samples of the task. Similarly, the approach in [22] also encodes the nominal behavior including the expected force profile in a probabilistic model. In comparison, we only require a single reference sample to distinguish contact events from normal execution, which speeds up the task definition in real systems. Further, we do not rely on additional sensors such as external vision systems or sensor arrays since the contact point is irrelevant to our system. The use of two Recurrent Neural Networks and a contact point estimation in [21] introduces a delay between contact and reaction of about 1 s, which hinders a quick reaction for collisions. We tackle this problem by the usage of a Support Vector Machine for fast online prediction. Similar to other data-driven methods, our approach does not

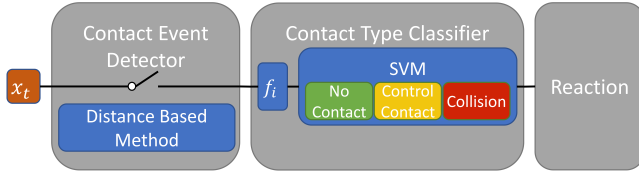


Fig. 3. Contact event pipeline for the detection, classification and reaction to contact events in robotic scenarios, where f_i are the extracted features.

need manually tuned contact force thresholds since these are chosen automatically through a changepoint detection algorithm in the training phase.

III. METHOD

In this section, we describe a system to distinguish task-contact from human-interaction and collision and to classify the intent of the interaction. For this purpose, we introduce a contact event pipeline that consists of a contact event detector (CED) and a contact type classifier (CTC) in order to facilitate reaction strategies.

A. Technical Approach

Learning from Demonstration (LfD) via kinesthetic teaching is used to transfer a task to the robot in a single trial. This avoids the effort of manual programming of the task and makes the procedure intuitive and open to non-expert users. A task is executed with a manipulator that has a dynamic model of the form

$$\tau_{\text{joint}} = M(q)\ddot{q} + c(q, \dot{q}) + g(q) + h(q, \dot{q}) + \tau_{\text{ext}}. \quad (1)$$

Here, $M(q)$ is the symmetric and positive-definite inertia matrix, $c(q, \dot{q})$ denotes the Coriolis and centrifugal torque, $g(q)$ describes the gravitationally induced torques and $h(q, \dot{q})$ includes further torques resulting from damping and friction. Furthermore, the joint torques τ_{joint} are influenced by external torques τ_{ext} resulting from external loads, such as contacts with the environment or the user. An estimate of the external joint torques $\hat{\tau}_{\text{ext}}$ can be derived from

$$\hat{\tau}_{\text{ext}} = \tau_{\text{joint}} - \tilde{M}(q)\ddot{q} - \tilde{c}(q, \dot{q}) - \tilde{g}(q) - \tilde{h}(q, \dot{q}), \quad (2)$$

where the friction has to be modeled to derive an accurate estimate of $\hat{h}(q, \dot{q})$. The parameters of \tilde{M} , \tilde{c} , \tilde{g} are the internal model of the corresponding mechanical terms in (1). Such an estimate of $\hat{\tau}_{\text{ext}}$ is already available on some robot platforms. An estimate of the forces and torques exerted on the external force/torque sensor can be derived by compensating the gravitational influences of the tool. This process is generally known as force/torque compensation as described in [23].

We apply our classification scheme on the execution phase of the task, where most harm can come to the human. The entire process must therefore be monitored continuously and any harmful situation must be eradicated as quickly and efficiently as possible. We therefore introduce a contact event pipeline as depicted in Fig. 3. In the following, we introduce the main concepts of this pipeline.

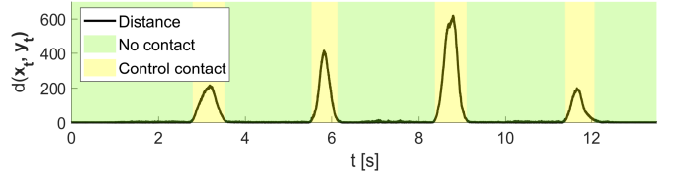


Fig. 4. Four contact events resulting from control contact indicated by comparing the current task execution to the reference sample with the distance metric. Phases of no contact and control contact as classified by the SVM indicated in green and yellow respectively.

B. Contact Detection

The first step in the contact event pipeline is the detection of a contact event, which is a simple boolean information, if at the current time instance a contact event occurred or is still ongoing. The contact event detector extracts all the sensor readings $\mathbf{x}_t = [p_t, \mathbf{o}_t, \boldsymbol{\tau}_t, \mathbf{w}_t, \mathbf{q}_t] \in \mathbb{R}^{27}$ that are characterized as being part of the contact event, where $p_t = [x, y, z] \in \mathbb{R}^3$ is the Cartesian end effector position, $\mathbf{o}_t = [q_w, q_x, q_y, q_z] \in \mathbb{R}^4$ the orientation in unit quaternions, $\boldsymbol{\tau}_t = [\tau_1, \dots, \tau_7] \in \mathbb{R}^7$ the external torques in each joint, $\mathbf{w}_t = [f_x, f_y, f_z, t_x, t_y, t_z] \in \mathbb{R}^6$ the wrench from the external sensor and $\mathbf{q}_t = [q_1, \dots, q_7] \in \mathbb{R}^7$ the joint angles at each time step t .

To detect a possible contact event, we introduce a Distance-Based Method (DBM). This approach uses the task knowledge in the form of a reference sample obtained from the human demonstration. Deviations from the expected sensory data, resulting from human interaction, are detected as shown in Fig. 4. All internal and external sensor measurements are compared in each instance to the existing time step in the reference sample of the same task. Deviations that exceed a threshold are considered as contact. These deviations are found by computing a distance metric over the reference sample $\mathbf{y}_t = [p_t, \mathbf{o}_t, \boldsymbol{\tau}_t, \mathbf{w}_t, \mathbf{q}_t]$ and the executed recording \mathbf{x}_t . The distance metric measures the dissimilarity between \mathbf{x}_t and \mathbf{y}_t as a sum over all dimensions k , given as

$$d(\mathbf{x}_t, \mathbf{y}_t) = \sum_{k=1}^n (x_{t_k} - y_{t_k})^{1.5}, \quad n = \dim(\mathbf{x}_t). \quad (3)$$

This distance metric has been chosen after a preliminary comparison with the Mahalanobis distance, several forms of the Minkowski distance (L_p -norm) and mean square error. It is realtime capable since it does not rely on a time window, it can be quickly solved online, and it delivers an accurate representation of the dissimilarity of the samples given only a single reference trajectory. Small deviations resulting from sensor noise or environmental uncertainties are not exponentiated compared to formulations with a higher power. This method considers all sensor modalities of \mathbf{x}_t in which changes, induced by human interaction, can occur. Compared to methods [21] and [19] where only joint speeds, external torques or wrench are considered, our proposed method considers all measured modalities simultaneously to detect all possible kinds of contact. Offsets resulting solely from an error in the robot pose are neglected in the mentioned works. Therefore, even a single reference sample for the nominal measurements delivers accurate results. Furthermore, our

metric does not cause a further delay like classifier based methods but instantly indicates contact events due to its computational efficiency.

To make this approach fully data-driven, an algorithm is used to automatically find the start and end point of contact events to avoid manual labeling during the generation of the training data. Contact events found by this preprocessing step are then fed to the model in the training phase. This method guarantees a consistent labelling process since it does not depend on subjective means, such as hand selecting contact events from the given data profile as done in related works. To find contact events with the distance metric over the whole length of the sample T , a changepoint detection algorithm is used which finds rapid changes in a signal [24], [25]. It partitions the signal into sections by minimizing the total residual error between the signal and a predefined statistical property. The residual error is given by

$$J(K) = \sum_{r=0}^{K-1} \sum_{i=k_r}^{k_{r+1}-1} \Delta(d_i; \chi([x_{k_r}, \dots, x_{k_{r+1}-1}])) + \beta K, \quad (4)$$

where k_0 and k_K are the first and the last sample of the signal, χ denotes the empirical estimate and Δ is the measured deviation. A penalty is added for an increasing number of changepoints K and by the proportionality factor β . The empirical estimate can be one of a statistical property. A comparison of possible statistical properties has proven that the root-mean-square level delivers the most accurate results in finding the start and end of a contact event even with multiple contact events in one sample recording. The root-mean-square level uses a constant mean set to zero and a piecewise constant variance. By maximizing the log-likelihood with respect to the variance and a fixed mean [24], equation (4) reduces to

$$\sum_{i=k_r}^{k_{r+1}-1} \Delta(d_i; \chi([x_{k_r}, \dots, x_{k_{r+1}-1}])) = (k_{r+1} - k_r) \log \left(\frac{1}{k_{r+1} - k_r} \sum_{i=k_r}^{k_{r+1}-1} x_r^2 \right). \quad (5)$$

The number of changepoints can be specified a-priori. The penalty term is therefore statically increased towards K , minimizing the risk of too many changepoints. From the findings of this algorithm, a threshold is derived for online contact event detection. This threshold is based on the mean μ and standard deviation σ of the distance at the changepoints of the training set with a factor k to increase robustness. A contact event occurs if

$$d(\mathbf{x}_t, \mathbf{y}_t) > \mu + k \sigma. \quad (6)$$

C. Contact Classification

In the second step of the contact event pipeline, the contact type classifier (CTC) uses only the data samples characterized as to be part of the contact event to classify the contact type. The CTC predicts one of the three different classes: “no contact”, “control contact” and “collision”. Contacts resulting from task-contact are already modeled

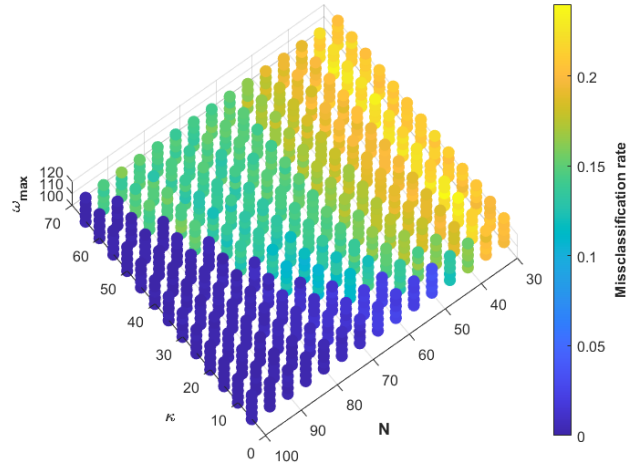


Fig. 5. K-fold cross validation of parameters N , ω_{\max} and κ .

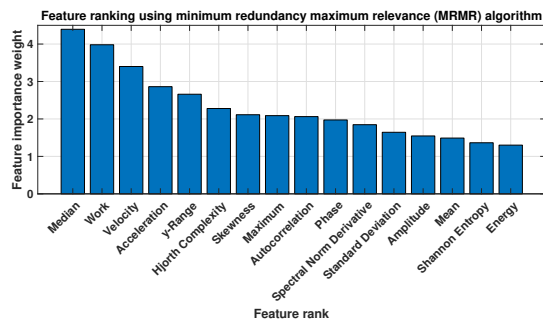
by the reference sample and are therefore ignored by the contact detector. The class of “no contact” is included to filter false positives indicated by the CED, which can result from high sensor noise. The classification is performed based on features derived from a sliding window of the instances in $\mathbf{X}_{t, \dots, t+N} = [\mathbf{x}_t, \dots, \mathbf{x}_{t+N}] \in \mathbb{R}^{N \times 26}$, where N is the length of the sliding window. A trade-off for N must be found between accuracy and delay. The length of N directly factors in to the delay $t_{\text{delay}} = N/f_s$ between contact and reaction. We use a sample frequency f_s of 500 Hz. On the other hand, the accuracy increases with the number of evaluated samples N . The optimal value is found through k-fold cross validation of the training set with different numbers for N , delay κ for the autocorrelation and the upper frequency bound ω_{\max} of the feature calculation, depicted in Fig. 5. The lowest missclassification rate is reached at $N = 85$ instances ($\hat{=} 0.17$ s).

The features computed from the sliding window N are both linear and non-linear as suggested in [12]. By computing feature values from all variables instead of using the measurements directly, we eliminate the need to know the point of contact and the direction of the contact force vector. Furthermore, these features give insights on the contact physics, which is beneficial compared to black box classification if the approach is applied in a real industrial environment and high transparency is desired. The predictors are classified by a non-linear, multi-class SVM. The main reasons for choosing a SVM are that they are among the fastest classifiers in the field and perform well with less training samples compared to the number of features. Furthermore, they excel tolerating irrelevant and redundant attributes [26].

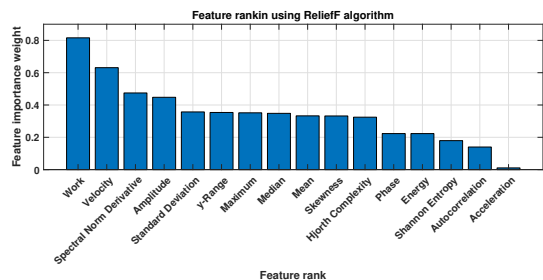
The Hjorth Complexity along other features is a widely used feature in touch modality identification via tactile sensing [27], [28]. When it comes to distinguish between collision and control contact, an important aspect can be the motion dynamics. Therefore, features of velocity and acceleration are included. Kouris et al. [13] draw the line

between collision and control contact based on the rate of change of a predefined frequency band. The Spectral Norm Derivative is the time derivative of the 1-norm of a predefined frequency band defined by an upper and lower bound. We implement a frequency band in the positive spectrum which is defined only by an upper bound ω_{\max} . Additionally, the features about virtual work (f_{16}, f_{17}) indicate if and how much energy is transferred from the robot to the environment or vice versa. A complete list of all derived features is presented in Table I.

The effectiveness of the selected features is evaluated with two algorithms: The ReliefF [29] and the Minimum Redundancy Maximum Relevance (MRMR) algorithm [30]. With the MRMR we analyze if any of the chosen features is irrelevant. Additionally, the robustness of the features is analyzed by the ReliefF algorithm, which uses a nearest neighbor type approach. Predictor weights are computed by rewarding if the k-nearest neighbors of a certain observation are within the same class. The evaluation results of the computed predictor data is depicted as feature rankings in Fig. 6(a) and Fig. 6(b) respectively. For the MRMR algorithm, a small drop in the weights indicates that the difference in the predictor importance is not significant. According to the ReliefF algorithm, only the acceleration features show a low level of robustness in the observations. By removing any of the less robust features, the model accuracy does not significantly increase. The same accounts for a feature reduction with a principal component analysis.



(a) MRMR algorithm for feature ranking.



(b) ReliefF algorithm for feature ranking.

Fig. 6. Results from the MRMR algorithm to determine relevance and redundancy and the ReliefF algorithm to determine the robustness of the chosen features.

D. Possible Reaction Strategies

During nominal task execution without human interaction at which both movements in free air and task-contacts are

TABLE I
LIST OF ALL FEATURES, WITH MATHEMATICAL FORMULATION, USED FOR CONTACT CLASSIFICATION.

Feature List		
f_1	Mean	$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$
f_2	Standard Deviation	$\sigma = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$
f_3	Maximum	$\max(x)$
f_4	y-Range	$\max(x) - \min(x)$
f_5	Median	$\tilde{x} = \text{median}(x)$
f_6	Hjorth Complexity	$C(x) = \frac{\frac{\sigma_{dd}}{\sigma_d}}{\frac{\sigma_d}{\sigma_x}}$
f_7	Shannon Entropy	$H(x) = - \sum_{i=0}^m p(x_i)^2 \log(p(x_i))^2$
f_8	Energy	$E_s = \int_{-\inf}^{\inf} \ x_t\ ^2 dt$
f_9	Skewness	$s = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})}{\sigma^3}$
f_{10}	Autocorrelation	$a = \frac{\sum_{i=1}^{N-\kappa} (x_i - \bar{x})(x_{i+\kappa} - \bar{x})}{\sum_{i=1}^{N-\tau} (x_i - \bar{x})^2}$
f_{11}	Phase	$\text{Phase}(\text{FFT}(x)) = \arctan \frac{\text{Im}\{\text{FFT}(x)\}}{\text{Re}\{\text{FFT}(x)\}}$
f_{12}	Amplitude	$\frac{1}{N} \sqrt{[\text{Re}\{\text{FFT}(x)\}]^2 + [\text{Im}\{\text{FFT}(x)\}]^2}$
f_{13}	Velocity	$v = \left\ \frac{\partial(x)}{\partial t} \right\ _2$
f_{14}	Acceleration	$a = \left\ \frac{\partial(v)}{\partial t} \right\ _2$
f_{15}	SND	$I_{\text{SND}} = \frac{\partial \ f_{\omega_{\min}:\omega_{\max}}\ _1}{\partial t}$
f_{16}	Work in Joints	$W_J = \sum_{t=t_1}^{t_2} \frac{\partial A_t}{\partial t} T_t^T$
f_{17}	Work in Sensor	$W_S = \sum_{t=t_1}^{t_2} \frac{\partial [P_t, O_t]}{\partial t} W_t^T$

possible, impedance control is advised. It is most common for compliant in contact applications where a manipulator interacts with the environment under uncertainties.

1) *Control contact*: If the user applies control contact, for instance to reshape the robot trajectory or to give a predefined touch command to the robot, an increasing force occurs between end effector state and commanded robot state. This can be encountered by a compliant robot behavior such as admittance control, e.g. as applied in [21].

2) *Collision*: In case of a collision, the main concern is to restore a safe state for the user as efficient and quick as possible. This reaction must cover both dynamic interactions as well as clamping situations. A "bounce back" in the direction of the most relief is a possibility [31]. Such a momentum observer-based reaction only requires proprioceptive sensor information and no external contact point estimation.

IV. EVALUATION

The viability of the contact event pipeline is tested both offline and in a real experimental task. First, it is elaborated on how training samples are generated and how the models are trained. More specifically, the accuracy of contact detection as well as the prediction accuracy of the correct contact type by the model, are tested. An impedance controlled KUKA LWR-IV+ robot with a wrist mounted JR3 force/torque-sensor and a Weiss Robotics WSG50 gripper are used for both the data generation and the experiment.

	Collisions	Control Contacts
Number of recordings	12	9
Total number of contact events	35	33
Types of sample motions	<ul style="list-style-type: none"> ▪ Collisions with arm ▪ Collisions with shoulder ▪ Collisions with head ▪ Impacts with robot arm and tool 	<ul style="list-style-type: none"> ▪ Linear and angular motions ▪ Control contacts on robot arm and tool ▪ Applied forces and torques

Fig. 7. Total samples and contact events recorded for model training and offline evaluation.

A. Training and Evaluation Samples

Initially, motion data is generated from kinesthetic teaching. The robot is put into gravity compensation and physically guided through the desired trajectory by the user. In total, 14 different types of movements have been recorded. Each movement consists of a unique set of motions to encapsulate a wide range of motions from real robotics tasks e.g. pick and place, pushing a button, moving etc. We ensured that different end effector orientations, various linear movements as well as different nullspace joint configurations are among the set. This allows the algorithm to learn from a variety of forces and torques applied on any joint. Once the set of representative types of movements is recorded, they are executed once on the robot without human interaction. These samples are later used as reference samples.

For control contacts, forces and torques are applied both on the entire robot surface as well as at the tool while the robot performs linear and angular motions simulating real world tasks. The contacts are of large variety to teach the model with a universal set of interactions. Collision samples are gathered using a human like dummy, shown in Fig. 8.

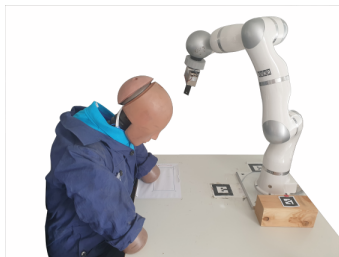


Fig. 8. Dummy for collision sample recording.

Five types of movements are chosen to impact the dummy on the arm, shoulder and head. This represents the human body parts that would most likely be affected in a real collision scenario with a single arm robot. Clamping situations are included. Each run of the training samples contains multiple contact

instances. All recorded types of motions for every contact class are summarized in Fig. 7.

Note that each contact event consists of multiple time steps (sample rate 2 ms) and that collisions are slightly shorter than control contacts. In summary, the model is trained on 14393 time samples which results in 28.79 s of human contact events.

B. Offline Evaluation

The offline evaluation is conducted, on a part of the initial sample set not part of the training set, to prove that the method is able to generate valid results for unknown data before users interact with the robot. The main focus

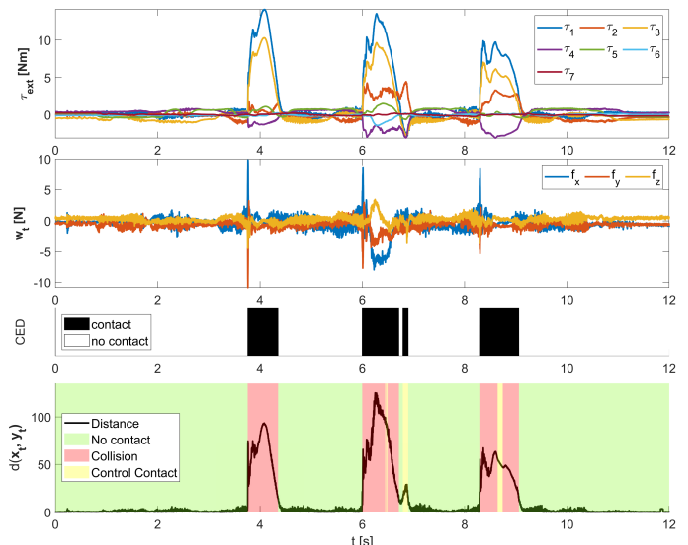


Fig. 9. Offline evaluation of three collisions. The first two plots show the torque and force profile from the proprioceptive and the external sensors. Row three is the output of the CED. The offset between the reference sample and the sample including human interaction is indicated as distance plot along the predicted contact class in the lowest plot.

points are: (i) how accurate and quick human interactions are determined from the data, (ii) the accuracy of the contact type prediction, especially for the initial instances of the contact occurrence, and (iii) filtering performance for false positive contacts.

1) *Results:* In the following evaluation, the parameters chosen according to the optimization are $N = 85$, $\omega_{\max} = 110$, $\kappa = 10$ and $k = 8$. Figure 9 clearly shows three simulated collisions as indicated by the CED. The data underlines, that a high rate of change in the contact forces indicate a quick change of momentum, which is typical for a collision. The reader's attention should be directed towards the initially classified instances of the individual contact events, which are correctly classified (red area). This section is of the most importance, since a possible reaction is based on this early prediction. In Fig. 4, four control contacts are analyzed, which are classified entirely correctly.

The prior human interactions occurred with the robot in free motion. In Fig. 10, the robot is in the state of task-contact from $t \approx 3.0$ s to $t \approx 5.2$ s and from $t \approx 5.8$ s to $t \approx 8.8$ s where an object is slid along a horizontal surface. During the task-contact, two control contact events are performed ($t \approx 3.9$ to $t \approx 4.9$ and $t \approx 6.9$ to $t \approx 7.8$). Both interactions are initially classified correctly but show misclassification due to the low compliance of the robot. The impact, resulting of the tool non-smoothly reconnecting with the object, ($t \approx 4.9$ s and $t \approx 7.8$ s) is also classified as collision. Another important aspect of this result is the filtering capability of the model. False positives at the beginning of ($t \approx 3.0$ s and $t \approx 5.8$ s) and during the task-contact ($t \approx 8.6$ s) are almost entirely filtered out. The overall accuracy for contact event detection in the offline evaluation is at 99.6%. The remaining, single instances, are not numerous enough to trigger a reaction.

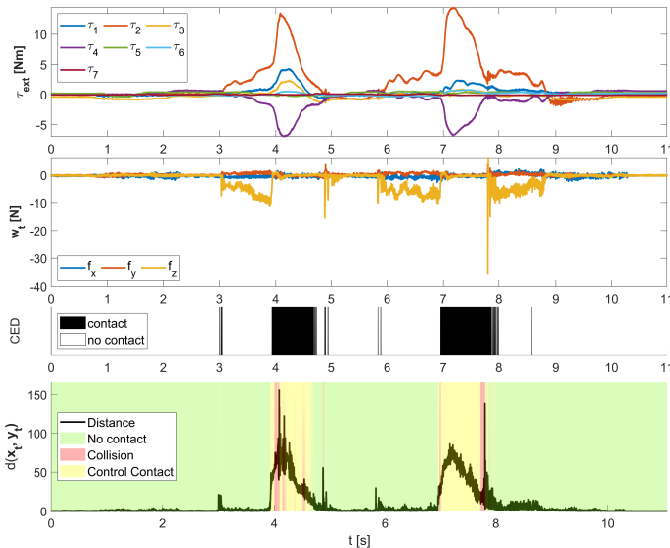


Fig. 10. Offline evaluation of two control contacts during an occurring task-contact. The first two plots show the torque and force profile from the proprioceptive and the external sensors. Row three is the output of the CED. The offset between the reference sample and the sample including human interaction is indicated as distance plot along the predicted contact class in the lowest plot.

2) *Discussion:* To generate the data for offline evaluation, the dummy is set up to intersect the anticipated trajectory of the robot. After the first rather elastic collision, the dummy is further pushed by the robot which causes a misclassification in the ongoing contact state. Additionally, during the generation of the data, no reaction behavior is employed to examine the full duration of the contact event. The offline evaluation results are depicted in Fig. 11. The soundest interpretation of these results is a contact-wise evaluation. As discussed, the initial classified instances are considered for the reaction and therefore determine the whole contact event. By this metric, the offline classification results are evaluated, which has been also used in [19]. The results are depicted in Fig. 11(a). The simulated trials involve both human contact during free motion and during task-contact. We achieve an accuracy of 92.5% over all contact events with slightly less precision for control contacts compared to collisions. The later discussed falsely classified instances during collision state drastically impact the accuracy of the sample-wise evaluation of the classification results (Fig. 11(b)). Considering the fact that no reaction is applied during the data recording phase biases the sample-wise result. Therefore, the focus is put on the contact-wise evaluation.

C. Experiment

The goals of the experiment are to evaluate the realtime capability of the framework, the adaptability of the system to unknown settings and its performance in a real-world scenario. The setting of the experiment is chosen so that interactions with the robot can both be performed separately and during task contact. The scenario chosen is a paint roller task. A paint roller is attached to the end effector as depicted in Fig. 1. From its initial position, the robot moves towards the table until the paint roller touches the surface and applies

	Cont.	Coll.	Prec.
Cont.	24	3	88.9%
Coll.	1	18	94.7%
Rec.	96.0%	85.7%	92.5%

	Cont.	Coll.	Prec.
Cont.	14510	12344	54.0%
Coll.	780	5170	86.9%
Rec.	94.9%	70.5%	60.0%

(a) Confusion matrix for contact-wise evaluation (b) Confusion matrix for sample-wise evaluation

Fig. 11. Offline evaluation results as confusion matrices. Contact-wise evaluation (a) by considering first classified instances of each contact event and sample-wise evaluation (b) by considering all instances.

	Cont.	Coll.	Prec.
Cont.	18	1	94.7%
Coll.	2	19	90.5%
Rec.	90.0%	95.0%	92.5%

	Cont.	Coll.	Prec.
Cont.	70	15	82.4%
Coll.	10	65	86.7%
Rec.	87.5%	81.3%	84.4%

(a) Confusion matrix for trained user (b) Confusion matrix for novel users

Fig. 12. Experimental results as confusion matrices. Results of twenty contact events for each contact type per participant for a user that previously participated in generation of training samples (a) and for novel users (b).

small amounts of pressure. The contact to the surface is then kept and the robot performs a painting motion. Thereafter, the contact is discontinued and the arm moves to a second position where the motion is repeated. The user is informed of the current contact situation and type by a dialog window, indicating the type of contact as well as the progress of the current task.

1) *Results:* A small user study is conducted with one user priorly involved in the generation of the training samples and four novel users. The participants are asked to interact with the robot during the task as it would occur most natural and intuitive to them. The robot is interacted with during free motion and task-contact, on several positions on the robot surface including the tool and paint roller (Figures 1(a), 1(b), 1(d)). Collisions are initiated with the hand (Fig. 1(c)), arm and elbow and by inducing clamping of the human hand between the paint roller and the horizontal surface. For each type of contact, a total of twenty contacts are performed per user. A total of one hundred contacts are therefore recorded for each contact type. The outcome of the trials are depicted in Fig. 12 for the trained user (12(a)) and for the novel users (12(b)) separately.

2) *Discussion:* The results of the trained user show an accuracy of 92.5% which is equal to the offline evaluation, but with a higher and lower precision for control contact and collision respectively. The systems responds correctly to novel users with an accuracy of 84.4%. This is an 8.1% decrease compared to the trained user and is mostly due to lower recall value for collision. More cautious contacts (initiated collisions) could be an explanation for this result since a human naturally tries to avoid harmful situations to the own body. This does not apply to the trained user, where the accuracy is equal to offline evaluation. However, the precision is shifted. A significant higher precision is achieved for control contacts while the precision of collisions has dropped by 4.2%. This difference occurs because

of the further training of the user. Contacts are executed more cautious, shifting the focus towards control contact. The results are again obtained by a contact-wise evaluation metric. During the experiment we obtained a very low false positive rate for contact event detection. The accuracy lies at 99.9%.

V. CONCLUSION

In this work, a novel contact event pipeline for human-robot interaction is proposed. It is shown that the system is able to actively monitor the execution of learned tasks and to distinguish human interaction and collision from task-contact. By classifying the interaction to determine the intent of the user, the system enables safe and intuitive human robot collaboration. More specifically, the set-up procedure lets the user easily record new tasks and monitor the execution without the need for time consuming manual programming. The proposed monitoring system can detect and classify contact events considerably quicker than comparable methods, enabling a fast and robust robot reaction. Based on only one reference sample, the system is able to detect contact events with increased robustness by filtering out false positives. The method is based only on proprioceptive sensor data and a wrist mounted force/torque sensor, therefore avoiding additional time consuming and costly sensor data evaluation. We have proven, by way of experiment and user study, the adaptability of the system to unknown scenarios and the intuitiveness and safety gained during human-robot collaboration.

REFERENCES

- [1] "Robots and the workplace of the future," *International Federation of Robotics*, 2018.
- [2] "Demystifying collaborative industrial robots," *International Federation of Robotics*, 2018.
- [3] J. Park, S. Haddadin, J. Song, and A. Albu-Schäffer, "Designing optimally safe robot surface properties for minimizing the stress characteristics of human-robot collisions," in *2011 IEEE International Conference on Robotics and Automation*, 2011, pp. 5413–5420.
- [4] S. Haddadin, S. Haddadin, A. Houry, T. Rokahr, S. Parusel, R. Burgkart, A. Bicchi, and A. Albu-Schäffer, "A truly safely moving robot has to know what injury it may cause," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2012, pp. 5406–5413.
- [5] Y. Yamada, Y. Hirasawa, S. Huang, Y. Umetani, and K. Suita, "Human-robot contact in the safeguarding space," *IEEE/ASME Transactions on Mechatronics*, vol. 2, no. 4, pp. 230–236, 1997.
- [6] K. Ikuta, H. Ishii, and M. Nokata, "Safety evaluation method of design and control for human-care robots," *The International Journal of Robotics Research*, vol. 22, no. 5, pp. 281–297, 2003.
- [7] M. Zinn, B. Roth, O. Khatib, and J. K. Salisbury, "A new actuation approach for human friendly robot design," *The International Journal of Robotics Research*, vol. 23, no. 4-5, pp. 379–398, 2004.
- [8] A. Bicchi and G. Tonietti, "Fast and "soft-arm" tactics [robot arm design]," *IEEE Robotics Automation Magazine*, vol. 11, no. 2, pp. 22–33, 2004.
- [9] D. M. Ebert and D. D. Henrich, "Safe human-robot-cooperation: Image-based collision detection for industrial robots," in *IEEE/RSJ international conference on intelligent robots and systems*, vol. 2. IEEE, 2002, pp. 1826–1831.
- [10] S. Haddadin, A. Albu-Schäffer, A. De Luca, and G. Hirzinger, "Collision detection and reaction: A contribution to safe physical human-robot interaction," in *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2008, pp. 3356–3363.
- [11] K. Hoang Dinh, O. S. Oguz, M. Elsayed, and D. Wollherr, "Adaptation and transfer of robot motion policies for close proximity human-robot interaction," *Frontiers in Robotics and AI*, vol. 6, p. 69, 2019.
- [12] S. Golz, C. Osendorfer, and S. Haddadin, "Using tactile sensation for learning contact knowledge: Discriminate collision from physical interaction," pp. 3788–3794, 2015.
- [13] A. Kouris, F. Dimeas, and N. Aspragathos, "A frequency domain approach for contact type distinction in huma-robot collaboration," *IEEE Robotics and Automation Letters*, vol. 3, no. 2, pp. 720–727, 2018.
- [14] A. Cherubini, R. Passama, A. Crosnier, A. Lasnier, and P. Fraise, "Collaborative manufacturing with physical human-robot interaction," *Robotics and Computer-Integrated Manufacturing*, vol. 40, pp. 1 – 13, 2016.
- [15] D. Lee and C. Ott, "Incremental kinesthetic teaching of motion primitives using the motion refinement tube," *Autonomous Robots*, vol. 31, no. 2, p. 115131, 2011.
- [16] C. Burghart, S. Yigit, O. Kerpa, D. Osswald, and H. Woern, "Concept for human robot co-operation integrating artificial haptic perception," in *Intelligent Autonomous Systems*, vol. 7, 2002, pp. 38–45.
- [17] G. Michalos, S. Makris, N. Papakostas, D. Mourtzis, and G. Chrysolouris, "Automotive assembly technologies review: challenges and outlook for a flexible and adaptive approach," *CIRP Journal of Manufacturing Science and Technology*, vol. 2, no. 2, pp. 81 – 91, 2010.
- [18] S. Haddadin, A. De Luca, and A. Albu-Schäffer, "Robot collisions: A survey on detection, isolation, and identification," *IEEE Transactions on Robotics*, vol. 33, no. 6, pp. 1292–1312, 2017.
- [19] N. Briquet-Kerestedjian, A. Wahrburg, M. Grossard, M. Makarov, and P. Rodriguez-Ayerbe, "Using neural networks for classifying human-robot contact situations," in *2019 18th European Control Conference (ECC)*. IEEE, 2019, pp. 3279–3285.
- [20] T. Eiband and D. Lee, "Identification of common force-based robot skills from the human and robot perspective," in *Humanoid Robots (Humanoids), IEEE-RAS 20th International Conference on*. IEEE, 2021.
- [21] M. Lippi and A. Marino, "Enabling physical human-robot collaboration through contact classification and reaction," in *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, pp. 1196–1203.
- [22] D. Park, Z. Erickson, T. Bhattacharjee, and C. C. Kemp, "Multimodal execution monitoring for anomaly detection during robot manipulation," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, 2016, pp. 407–414.
- [23] D. Kubus, T. Kroger, and F. M. Wahl, "On-line rigid object recognition and pose estimation based on inertial parameters," in *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2007, pp. 1402–1408.
- [24] M. Lavielle, "Using penalized contrasts for the change-point problem," *Signal Processing*, vol. 85, no. 8, pp. 1501 – 1510, 2005.
- [25] R. Killick, P. Fearnhead, and I. A. Eckley, "Optimal detection of changepoints with a linear computational cost," *Journal of the American Statistical Association*, vol. 107, no. 500, pp. 1590–1598, 2012.
- [26] H. Bhavsar and A. Ganatra, "A comparative study of training algorithms for supervised machine learning," *International Journal of Soft Computing and Engineering (IJSCE)*, vol. 2, 01 2012.
- [27] M. Kaboli, A. Long, and G. Cheng, "Humanoids learn touch modalities identification via multi-modal robotic skin and robust tactile descriptors," *Advanced Robotics*, vol. 29, no. 21, pp. 1411–1425, 2015.
- [28] T. Balli Altuglu and K. Altun, "Recognizing touch gestures for social human-robot interaction," in *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*, ser. ICMI '15. New York, NY, USA: Association for Computing Machinery, 2015, pp. 407–413.
- [29] M. Robnik-Šikonja and I. Kononenko, "Theoretical and empirical analysis of relieff and rrelieff," *Machine learning*, vol. 53, no. 1-2, pp. 23–69, 2003.
- [30] C. Ding and H. Peng, "Minimum redundancy feature selection from microarray gene expression data," *Journal of bioinformatics and computational biology*, vol. 3, no. 02, pp. 185–205, 2005.
- [31] A. De Luca, A. Albu-Schäffer, S. Haddadin, and G. Hirzinger, "Collision detection and safe reaction with the dlr-iii lightweight manipulator arm," in *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2006, pp. 1623–1630.