Supervised learning of gesture-action associations for human-robot collaboration

Dadhichi Shukla, Özgür Erkent, Justus Piater
Intelligent and Interactive Systems, University of Innsbruck, Austria
firstname.lastname@uibk.ac.at

***Abstract*— As human-robot collaboration methodologies develop robots need to adapt fast learning methods in domestic scenarios. The paper presents a novel approach to learn associations between the human hand gestures and the robot's manipulation actions. The role of the robot is to operate as an assistant to the user. In this context we propose a supervised learning framework to explore the gesture-action space for human-robot collaboration scenario. The framework enables the robot to learn the gesture-action associations on the fly while performing the task with the user; an example of zero-shot learning. We discuss the effect of an accurate gesture detection in performing the task. The accuracy of the gesture detection system directly accounts for the amount of effort put by the user and the number of actions performed by the robot.**

I. INTRODUCTION

One of the major challenges for a robot in human-robot interaction (HRI) scenarios is to explore the large state-space of the environment. To perform a manipulation action a robot should be aware of the three main states: the state of the human, the state of the objects to manipulate, and the robot's own state. The state of the human is the command given by the user which in this work is a static hand gesture. The state of an object is whether it is in the robot's hand or not.

In this paper we propose a novel framework to learn associations between the hand gestures and the robot actions. Although the user sees the associations between gestures and actions the robot actually learns associations between the state of the system (i.e. the state of the human, the state of the object, and the robot's own state) and the action performed. Initially the associations are unknown which are learnt in a supervised zero-shot learning fashion on the fly.

We demonstrate our framework in a domestic scenario to assemble furniture e.g. a table, as illustrated in Fig. 1. The robot assists the user in the assembly of the table. Its objective is to handover the table legs to the user. We define 10 ways to handover legs. In order to evaluate the proposed framework extensively we assemble 5 tables 20 times. It would be tedious and practically expensive to perform such a high number of table assemblies in a real environment. Therefore experiments are conducted in a simulated environment.

A. Motivation and Contribution

Numerous studies [1], [2] have reported methods for robots to learn to complete tasks more efficiently i.e. completion in the least number of actions. Our work is motivated by the Joint Intention Theory by Cohen et al. [3] where authors propose a formal approach to building artificial collaborative agents. The authors describe *collaboration* not only as the commitment of members to achieve the goal but also–if necessary–having a mutual belief about the state of the goal.

978-1-5090-4023-0/17/$31.00 ©2017 IEEE

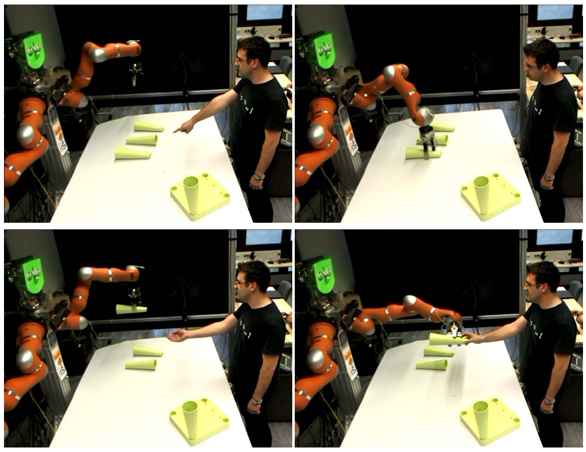

Fig. 1: Robin (the robot) assists the user in the assembly of a table. The user performs gestures like 'pointing' and 'give me' to which the robot reacts as 'grasp object' and 'hand over', respectively.

The furniture assembly task takes place in a close proximity hence it is irrelevant to observe full-body human pose. The gestures are performed with hand and fingers, not with body and arm [4]. Therefore, we choose to interact using the hand gestures since they inherently provide spatial information of the user's hand. For example, gestures like pointing can be used to localize objects [4]. Moreover, the hand gestures are comparatively natural than a computer interface. We use gestures from the Innsbruck Multi-view Hand Gestures (IMHG) dataset [5]. The gestures in the IMHG dataset are closely related to the semantic content of verbal language and were designed based on the HRI study conducted by Jensen et al. [6].

The approach is statistically driven, in other words, every data point is an input-output pair where the input maps to an output. This grants the flexibility to implement the Proactive Incremental Learning (PIL) framework. The robot receives a feedback (positive or negative) based on the action it has performed with respect to the state of the system. The gesture-action associations are recorded on receiving the feedback consequently learning the mapping between the input i.e. the state of the system and the output i.e. the robot action.

It is a supervised learning approach, however, the *training* and the *testing* are not the two distinct phases of the process. Since associations are learnt on the fly both phases are active till the system reaches the goal. It records each data point incrementally to develop the gesture-action model. This *incremental* attribute of the PIL framework provides the freedom to the user to establish the gesture-action associations at will. The probabilities of the associations i.e. to perform an action given the state, are learnt as the task advances.

Recent HRI study by Jenson et al. [7] shows that people expect robot to act proactively with the same interactional levels as that of humans. The *proactive* nature of the PIL framework addresses this aspect. The robot is able to decide the most likely action to perform given the state of the system after learning the associations during the task. Additionally, the robot based on the acquired knowledge can correct the state of the detected hand gesture if it invokes an invalid state of the system. For example, a robot trying to grasp another object while an object is already in its hand.

The invalid state can occur due to two main reasons: (1) a gesture is detected with a low confidence score irrespective of the state of the system therefore it is discarded, (2) a gesture is misclassified and it is incompatible with the state of the system irrespective of the confidence score. The state of the human as understood by the robot depends on the rate of detection of the communicative signal. During an invalid state the robot looks at the learnt probabilities to correct the detected gesture. The number of interactions taking place during a human-robot collaboration task is directly proportional to the accuracy of the hand gesture detection system. We quantitatively study and discuss the effect of the rate of detection of gesture in section IV.

The main contributions of this study are summarized as

1) A supervised learning approach to design a framework– Proactive Incremental Learning (PIL) framework– to learn associations between the hand gestures and the robot's manipulation action on the fly,
2) Proactively correct the misclassified gesture,
3) Study the effect of the rate of detection of the hand gestures.

B. Related work

The human-robot interaction studies with an active human-in-loop involvement has been presented in many instances such as learning by imitation [8], demonstration [9], and feedback [10]. Interactive reinforcement learning (IRL) based approaches [1], [11] have demonstrated that human-generated reward can be powerful and is a fast learning technique over classic reinforcement learning.

Thomaz et al. introduced IRL which enables the user to provide positive and negative rewards using a computer interface in response to the manipulation action of the robot. These rewards are a guidance input which leads the robot to perform the desired behaviour [1]. Human-Robot Interaction Operating System (HRI/OS) designed by Fong et al. allow humans and robots to collaborate in joint tasks in a peer-to-peer fashion [12]. The key feature of their system is that the coordination takes place through a dialogue only when the help is asked.

Lutkebohle et al. proposed a bottom-up strategy where the robot performs an action then engages in a dialogue with the user [13]. Their framework focusses on two objectives: (1) What the robot has to learn, and (2) Bring attention of the user to provide feedback based on the action. Consequently, the robot develops associations between the verbal commands and actions during the training phase.

A commonality among the state of the art is there are two distinct phases for the robot viz., a learning phase and a testing phase. To learn new associations the system needs to be retrained which is not the case for the proposed work; an advantage of the PIL framework. An exception among the previous works is the approach by Grizou et al. [14] where the robot learns to interpret voice commands while learning a task. Their framework is based on inverse reinforcement learning where the user instructs the robot what to do.

The proposed PIL framework is guided by similar principles particularly the idea of feedback from the user. Like Grizou et al.'s framework the PIL also incorporates both the training and the testing during the task. A pertinent difference with the PIL framework is in our work the robot explores manipulation actions and is not specifically guided by the user regarding which action to perform. In addition to learning the associations on the fly these associations can be shared, modified, and updated with new associations while performing a new task.

II. HUMAN-ROBOT COLLABORATION SCENARIO

A. Domestic Scenario

We designed the HRI setup with 'Robin' (the humanoid) which has 2 KUKA light weight robot arms, 2 Schunk hands, and a KIT robotic head [15] in a virtual environment. We use the left arm of the robot to interact since gestures in the IMHG dataset are commonly performed with the right hand.

We use 6 types of hand gestures from the IMHG dataset. These gestures are categorized based on Quek's taxonomy [16] as shown in Fig. 2. Let $G = \{\texttt{pointing}, \texttt{give me}, \texttt{grasp}, \texttt{release}\}$ be the set of 4 instructional gestures shown as *pointing*, *give me*, *grasp close*, and *grasp open*, respectively, from the dataset. The *pointing* gesture is used to indicate an object or a location in the workspace. The *fist* gesture is used to instruct the robot to pause. Since the setup is simulated the *fist* gesture is not included in this work. Let $F = \{\texttt{OK}, \neg\texttt{OK}\}$ be the set of feedback signals given by the user by performing *thumb up* and *thumb down* gestures, respectively. The feedback is in the form of a binary approval signal which determines how a gesture-action association is scored.

The 5 manipulation actions known to the robot are: open robot's hand (`open`), close robot's hand (`close`), go on the top of the pointed object (`object`), go towards the human hand (`human`), and go to the pointed location on the table (`location`). Let $A = \{\texttt{open}, \texttt{close}, \texttt{object}, \texttt{human}, \texttt{location}\}$ be the set of those actions. Since the state of the object mainly depends on

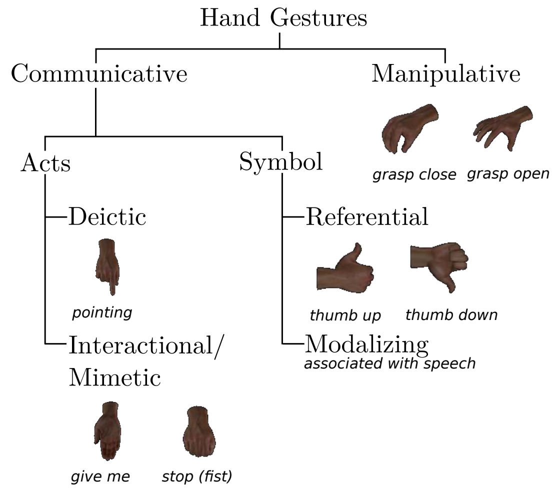

Fig. 2: Taxonomy of hand gestures.

Step t	Detected gesture g_t	Previous robot action a_t	Robot's hand h_t	Action a_{t+1}	Feedback η_{t+1}
1	pointing	open	free	close	¬OK
2		open	free	open	¬OK
3		open	free	object	OK
4	grasp	object	free	open	¬OK
5		object	free	close	OK
6	give me	close	occupied	location	¬OK
7		close	occupied	human	OK
8	release	human	occupied	object	¬OK
9		human	occupied	open	OK
10	pointing	open	free	object	OK
11	grasp	object	free	close	OK
12	pointing	close	occupied	close	¬OK
13		close	occupied	human	¬OK
14		close	occupied	location	OK
15	release	location	occupied	open	OK
16	pointing	open	free	object	OK
17	pointing	object	free	open	¬OK
18		object	free	object	OK
19	grasp	object	free	close	OK
20	give me	close	occupied	human	OK
21	pointing	human	occupied	object	¬OK
22		human	occupied	location	OK
23	release	location	occupied	open	OK

TABLE I: An example to handover legs of the table to learn the gesture-action associations.

the state of the robot's hand we define 2 states of the robot hand. The state of the robot's hand is considered free when the object is not in its hand (free) and it is considered as occupied when the object is in its hand (occupied). Let $H = \{\texttt{free}, \texttt{occupied}\}$ be the set of these states of the robot hand.

B. Task execution

The state s of the system at the time step t consists of three attributes given as $s_t = \langle g_t, a_t, h_t \rangle$, where $g_t \in G$ is the detected gesture, $a_t \in A$ is the state of the robot i.e. the action performed by the robot in the previous step, and $h_t \in H$ is the state of the robot's hand. At each step t of the assembly three entities are recorded: the state of the system s_t, the action robot will perform $a_{t+1} \in A$, and the feedback signal $\eta_{t+1} \in F$ given by the user after the action.

The gesture-action associations are mainly seen from the user's perspective while the state-action associations is what the robot learns. Henceforth, the state-action association and the gesture-action association are used interchangeably. Consider the sequence of quintuples as shown in Table I to handover legs of the table. Each block with the solid line is a type of a handover. For explanation purposes let us consider that gestures are detected accurately.

Initially, Robin is at a default position with its hand open and the four legs of the table are kept within its reachability. First, the user points at one of the leg. The vision system detects the hand gesture as pointing. Robin then selects one of the random actions from A, let's say, close i.e. close robot's hand. Since close was not the desired action the user gives a feedback signal ¬OK. Similarly, ¬OK is given to the next random action open. Finally, at step 3 when the robot performs the action object i.e. go on the top of the pointed object, the user gives OK.

Next, the user performs a grasp gesture to instruct the robot to grasp the object. Since Robin has never seen this state previously it randomly executes action open and in return the user gives feedback as ¬OK. It then performs action close and grasps the object to which a OK is signalled. In due time Robin would have learnt that it has to perform the action object given the state of the robot is open, its hand is free, and the user makes pointing gesture . Moreover, it also enables to design a intent prediction modal to speed up the task. For example, the user is most likely to perform grasp after pointing, therefore, the robot is aware that the action followed by going on the top of the object (object) will be most likely to grasp it (close).

III. Proactive Incremental Learning

The Proactive Incremental Learning (PIL) framework is an example of supervised zero-shot learning. It is designed to reach to the final state while minimizing the number of actions performed by the robot as well as the number of gestures shown by the user. Let's consider the table assembling task as described in section II-B. Let N be the number of legs which Robin has to handover to the user. The Proactive Incremental Learning framework consists of two modules: 1) Feedback based gesture-action associations, and 2) Proactive gesture correction.

A. Feedback based gesture-action associations

The goal of the system is to learn the probability of an action a_{t+1} to be executed by the robot given the state of the system s_t i.e. $P(a_{t+1}|s_t)$. Initially, the probabilities are distributed uniformly among all the robot actions. The robot incrementally learns the probabilities at every step when the user gives a feedback. The feedback is received as the binary score for the state-action association (s_t, a_{t+1}) as

$$\eta = \begin{cases} 1, & \text{feedback} = \texttt{OK} \\ 0, & \text{feedback} = \neg\texttt{OK}. \end{cases} \tag{1}$$

Let T be the 4-D table which stores the state-action associations at each step t. The score of each state-action association is recorded in cell $\mathsf{T}(s_t, a_{t+1})$. It is incremented based on the feedback η as

$$\mathsf{T}(s_t, a_{t+1}) = \begin{cases} \mathsf{T}(s_t, a_{t+1}) + \eta, & \eta = 1 \\ \mathsf{T}(s_t, a_{t+1}), & \eta = 0. \end{cases} \quad (2)$$

During the early steps of the task the robot performs random actions since it has not acquired any knowledge. Though in later steps it has updated the score for the state-action association (s_t, a_{t+1}) and it can compute $P(a_{t+1}|s_t)$. The probability of the action a given the state s of the system is computed as

$$P(a|s) = \frac{\mathsf{T}(s, a)}{\sum_{i=1}^{|A|} \mathsf{T}(s, a_i)}. \quad (3)$$

If the score values in T are normalized as per the joint probabilities then eq. 3 represents Bayes' rule. The best action a_{t+1} to perform given the state of the system is selected as

$$a^* = \underset{a_t}{\operatorname{argmax}}\, P(a_{t+1}|s_t). \quad (4)$$

In the human-robot interaction studies conducted by Jensen et al. [6], the authors observed that there is no feedback or rarely a positive feedback if the robot performs an expected action. Although users strongly give a negative feedback $\neg\texttt{OK}$ for an unexpected action. Therefore, in the PIL framework we consider *no feedback* too as `OK`.

Let us consider the example as described in Table I. It can be seen at step $t = 3$ the robot learns the association between the state of the system $s_3 = \langle\texttt{pointing}, \texttt{open}, \texttt{free}\rangle$ with the action $a_4 = \langle\texttt{object}\rangle$. At step $t = 10$ the probability $P(\texttt{object}|s_{10})$ has the highest probability. Consequently the robot chooses action `object` to perform instead of opting for a random selection.

In other instance at $t = 17$ from user's point of view it may seem that the robot will perform `object`. However, for the robot $s_{17} = \langle\texttt{pointing}, \texttt{object}, \texttt{free}\rangle$ which has never occurred previously. The probability of an action to perform given the state s_{17} is uniformly distributed among all the possible actions. Therefore, the probability $P(\texttt{object}|s_{17})$ is same as that of other actions and the robot opts to perform a random action.

B. Proactive gesture correction

The accurate detection of the gesture is vital in learning of the gesture-action associations. However, the gesture detection system may misclassify due to changes in the lighting conditions of the environment, differences in the appearance of a gesture among users, or a low confidence score in the detection, etc. To overcome this problem we incorporate a *gesture correction* module in the PIL framework.

The PIL framework enables the robot to correct the misclassified gesture based on the learnt gesture-action associations. The user is unaware if the system has misclassified the gesture. Though in case of misclassification if the detected gesture invokes a state which conflicts with the physical laws or which is redundant then the system can judge whether a gesture correction is needed or not.

A correction step can only be active after some associations are recorded in T. If an invalid state occurs during the task then the PIL framework looks at the sequence T to infer the most probably valid gesture. Let s_t be the invalid state of the system and therefore no action a_{t+1} is selected. At this point the system checks where has the state s_{t-1} occurred previously in T. The corrected state s'_t is most likely to be one of the states which occurs after one of the instances of s_{t-1} with feedback `OK`.

The robot chooses s'_t with the most occurrences i.e. with the highest probability, and corrects the state of the system s_t. It performs a'_{t+1} and it is recorded as a_{t+1}. For example, in Table I at $t = 21$ the user performs `poining` gesture. If it was incorrectly detected as `grasp` this would trigger an invalid state. The robot cannot grasp an object when the state of the robot hand is `occupied`. The system then decides that correcting the gesture is necessary and it looks for the gesture with highest probability which can occur after `give me`. The system corrects the gesture maximizing the probability $P(s'_t|s_{t-1})$. The detected gesture is corrected from `grasp` to `release` and the robot performs action `open` as the most likely action.

In the case of an invalid state it is possible that the corrected action was not the one that the user had expected. The user always has the freedom to give a $\neg\texttt{OK}$ feedback. The robot then selects the next best action. If none of the previously learnt actions receive a `OK` feedback for the given state then the robot explores the actions which had received $\neg\texttt{OK}$ feedback. A valid state-action association can be miscategorized as $\neg\texttt{OK}$ and an invalid state-action association can be miscategorized as `OK` if the gesture detection system has a poor accuracy. An advantage of the PIL framework is that at first it proactively decides to perform an action and if it is not the one user desires then it enables the user to choose the action as described in section III-A.

C. Algorithm

We initialize the number of objects N which are to be handed over to the user. The robot is set to a default position with its hand open and the step counter is initialized at $t = 1$. The pseudo-code of the PIL framework is as described in Alg. 1. An object is consider to be delivered to the user only if the robot had performed either `human` or `location` in the last step while holding the object and it performs `open` as the next action.

IV. RESULTS AND DISCUSSION

A. Simulation setup

We evaluate the PIL framework in a simulated environment. There are 4 instruction gestures, 2 feedback gestures, and 5 robot actions as described in section II-A. In this test scene we would like to assemble 10 four-legged tables i.e. in total the robot has to handover 40 table legs to the user. We define 10 ways of handover out of which 3 are as shown in Table I (step 1-9, 10-15, 16-23). In Table I user performs 14 instruction gestures to receive 3 objects. The number of command gestures g performed by the user in each handover

Algorithm 1: Proactive Incremental Learning framework

Initialize N, $a_t = \texttt{open}$, $h_t = \texttt{free}$, $\mathsf{T} = \emptyset$, $t = 1$
while $N > 0$ **do**
  Detect gesture g_t from the set of G
  Update $s_t = \langle g_t, a_t, h_t \rangle$
  if $\mathsf{T} \neq \emptyset$ **then**
    Compute a^* using eq. 4
    if (s_t, a_{t+1}) *is valid* **then**
      if a^* *is unique* **then**
        Perform action $a*$
        Update $\mathsf{T}(s_t, a_{t+1})$ as per eq. 2
      else
        Learn gesture-action associations as described in section III-A
    else
      Correct the state as described in section III-B
  else
    Learn gesture-action associations as described in section III-A
  if $(a_t = \textit{human} \vee a_t = \textit{location}) \wedge (h_t = \textit{occupied}) \wedge (a_{t+1} = \textit{open})$ **then**
    $N = N - 1$
    $t = t + 1$

ranges from 4-7. For the simulated environment we fed a predefined sequence of the gestures to the system to assemble 10 tables.

An important aspect in human-robot collaboration is the accurate detection of the command signal (here gestures). To evaluate the effect of accuracy we simulate the detection rate d of these gestures. For example, consider the sequence in Table I, if the detection rate of the system is $d = 60\%$ then only 8 randomly selected gestures will be detected correctly. It is to be noted that in order to simulate the feedback the correct state-action associations T_c are known to the system. A positive feedback is given only if the state-action association $(s_t, a_{t+1}) \in \mathsf{T}_\mathsf{c}$.

B. Simulation results

The goal of the framework is to reach the final state while minimizing the number hand gestures and the number of robot actions. In the PIL framework the robot is proactive which means the robot learns the associations on the fly, uses the knowledge to perform actions, and is able to correct the misclassified gestures. We simulate 4 detection rates $d = \{40\%, 60\%, 80\%, 100\%\}$.

The PIL framework is compared with a 2×2 pre-trained experimental setup at d detection rates. The two main axes of the pre-trained conditions are: (1) Trained at either 100% or d detection rate, and (2) System can or cannot perform gesture correction. The four conditions are

1) Both the training and the testing is conducted at d detection rate. It uses the learnt probabilities to correct the misclassified gesture.
2) The training is conducted at 100% detection rate and the testing at d detection rate. It uses the learnt probabilities to correct the misclassified gesture.
3) Both the training and the testing is conducted at d detection rate. In the case of incorrect gesture detection the user has to correct it manually i.e. to perform gesture repeatedly until it is correctly detected.
4) The training is conducted at 100% detection rate and the testing at d detection rate. In the case of correct gesture detection the user has to correct it manually i.e. to perform gesture repeatedly until it is correctly detected.

Since the training is on the fly in the PIL framework the comparison is done at d detection rate of the testing phase of the above conditions.

We assemble the set of 10 tables 20 times to obtain a statistically significant data. The comparison results of the 4 conditions with the PIL framework regarding the number of gestures and the number of robot actions required to complete the task are shown in Fig 3 and Fig 4, respectively. At a low gesture detection rate the number of gestures and the number of robot actions required for the PIL are higher with respect to condition 1 and condition 2. Due to a low detection rate a correct gesture-action association do not achieve highest probability as compared to the other possible gesture-action associations.

It is to be noted that the results include the number of robot actions and the number of hand gestures during the training period for conditions 1, 2, 3, and 4. The predefined sequence of gestures has 10 types of handovers. To cover all types of handovers at least twice the training of the aforementioned 4 conditions was done by assembling 5 tables. During the training it took $222, 203, 181, 144$ robot actions and $170, 148, 126, 111$ hand gestures to complete the task at $40\%, 60\%, 80\%, 100\%$ detection rate, respectively. An advantage of the PIL framework is that the user has the freedom to change gesture-action associations at any point in time; not possible with trained conditions.

We performed t-test on the data recorded from 20 iterations at each detection rate to check if the data points are significantly different from each other. The t-test is performed on the number of gestures and the number of robot actions required by the PIL framework and the 4 conditions. The p-values of our data indicate that the *null hypothesis* can be rejected with 5% significance level. Our simulation data is statistically significant with $p < 0.05$ having a maximum at $p_{\text{max}} = 0.048$.

V. CONCLUSION AND FUTURE WORK

We propose a supervised proactive incremental learning framework to learn the gesture-action associations on the fly. The learning phase in the PIL framework is active until the system reaches its final goal. Therefore, it is independent of an explicit training phase in comparison to previous approaches. In a real world human-robot collaboration scenario the communication signal is prone to changes in the environment

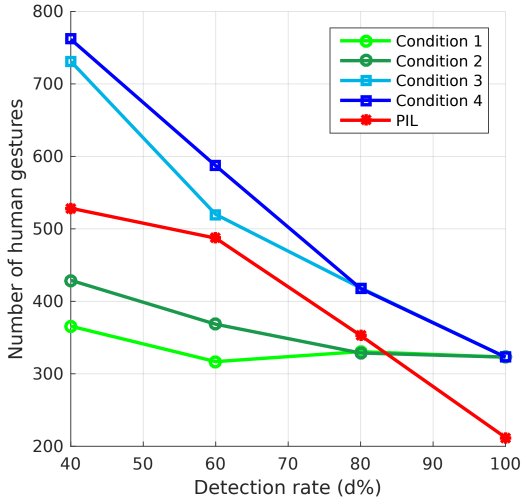

Fig. 3: Number of gestures performed by the user during the assembly of 10 tables with 4 objects.

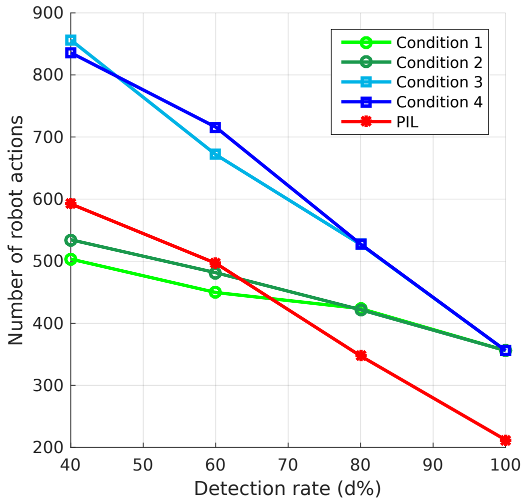

Fig. 4: Number of actions executed by the robot during the assembly of 10 tables with 4 objects.

or to the sensor noise. A poor signal-to-noise ratio can lead to misclassification of the instruction signal. This in turn can become tedious for the user. We studied the effect of the detection rate of the communicative signal on a table assembly task.

While the work proposed in this paper illustrates results of a simulated environment our next step is to test it with the real robot. We are working on techniques as discussed in [17] to use multiple cameras and learn the appearance of gestures on the fly in one-shot learning fashion. The PIL framework provides the flexibility to integrate a module for learning new associations. Additionally, the proactive nature will enable to design a action prediction modal. The robot will perform an action based on learnt probabilities instead of waiting for an instruction from the user.

VI. ACKNOWLEDGEMENTS

The research leading to this work has received funding from the European Community's Seventh Framework Programme FP7/2007-2013 (Specific Programme Cooperation, Theme 3, Information and Communication Technologies) under grant agreement no. 610878, 3rd HAND.

REFERENCES

[1] A. L. Thomaz and C. Breazeal, "Reinforcement learning with human teachers: Evidence of feedback and guidance with implications for learning performance," in *AAAI*, vol. 6, 2006, pp. 1000–1005.

[2] F. Cruz, S. Magg, C. Weber, and S. Wermter, "Training agents with interactive reinforcement learning and contextual affordances," *IEEE Transactions on Cognitive and Developmental Systems*, 2016.

[3] P. R. Cohen and H. J. Levesque, "Teamwork," *Nous*, vol. 25, no. 4, pp. 487–512, 1991.

[4] D. Shukla, O. Erkent, and J. Piater, "Probabilistic detection of pointing directions for human robot interaction," in *International Conference on Digital Image Computing: Techniques and Applications*, 11 2015. [Online]. Available: https://iis.uibk.ac.at/public/papers/Shukla-2015-DICTA.pdf

[5] ——, "A multi-view hand gesture rgb-d dataset for human-robot interaction scenarios," in *International Symposium on Robot and Human Interactive Communication*, 2016.

[6] L. Jensen, K. Fischer, D. Shukla, and J. Piater, "Negotiating Instruction Strategies during Robot Action Demonstration," in *10th ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 3 2015, pp. 143–144. [Online]. Available: https://iis.uibk.ac.at/public/papers/Jensen-2015-HRI.pdf

[7] L. Jensen, K. Fischer, F. Kirstein, D. Shukla, O. Erkent, and J. Piater, "It gets worse before it gets better," in *12th ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 3 2017. [Online]. Available: https://iis.uibk.ac.at/public/papers/Jensen-2017-HRI.pdf

[8] J. Bandera, J. Rodriguez, L. Molina-Tanco, and A. Bandera, "A survey of vision-based architectures for robot learning by imitation," *International Journal of Humanoid Robotics*, 2012.

[9] L. Rozo, P. Jiménez, and C. Torras, "A robot learning from demonstration framework to perform force-based manipulation tasks," *Intelligent service robotics*, 2013.

[10] W. B. Knox and P. Stone, "Reinforcement learning from human reward: Discounting in episodic tasks," in *RO-MAN, 2012 IEEE*, 2012.

[11] F. Cruz, J. Twiefel, S. Magg, C. Weber, and S. Wermter, "Interactive reinforcement learning through speech guidance in a domestic scenario," in *IEEE IJCNN, 2015*, 2015.

[12] T. Fong, C. Kunz, L. M. Hiatt, and M. Bugajska, "The human-robot interaction operating system," in *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*. ACM, 2006, pp. 41–48.

[13] I. Lutkebohle, J. Peltason, L. Schillingmann, B. Wrede, S. Wachsmuth, C. Elbrechter, and R. Haschke, "The curious robot-structuring interactive robot learning," in *Robotics and Automation, 2009. ICRA'09. IEEE International Conference on*. IEEE, 2009, pp. 4156–4162.

[14] J. Grizou, M. Lopes, and P.-Y. Oudeyer, "Robot learning simultaneously a task and how to interpret human instructions," in *2013 IEEE Third Joint International Conference on Development and Learning and Epigenetic Robotics (ICDL)*, 2013.

[15] T. Asfour, K. Welke, P. Azad, A. Ude, and R. Dillmann, "The Karlsruhe humanoid head," in *Humanoids 2008-8th IEEE-RAS International Conference on Humanoid Robots*. IEEE, 2008, pp. 447–453.

[16] F. K. Quek, "Eyes in the interface," *Image and vision computing*, vol. 13, no. 6, pp. 511–525, 1995.

[17] O. Erkent, D. Shukla, and J. Piater, "Integration of probabilistic pose estimates from multiple views," in *European Conference on Computer Vision*, 2016. [Online]. Available: https://iis.uibk.ac.at/public/papers/Erkent-2016-ECCV.pdf