COMMENTS TO SUBMIT – DRESSING

Details of setup

Robots

The dual-arm setup is composed of two Franka Emika Panda 7-DOFs[1] manipulators, whose kinematic chain is reported in Figure 1.

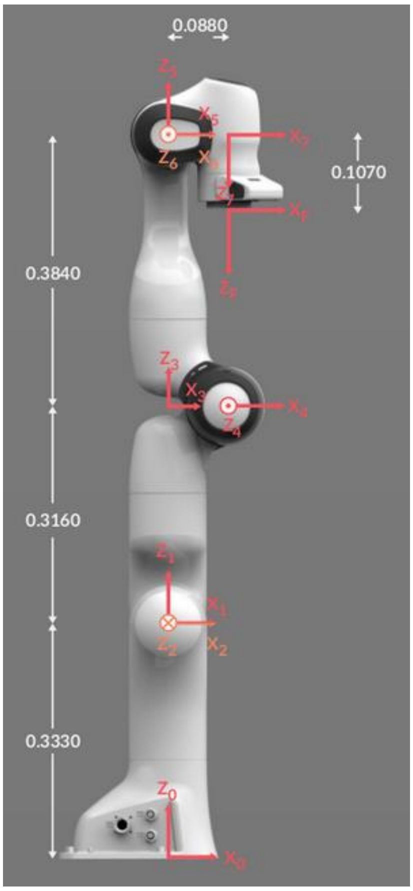

Figure 1 - Kinematic chain of Franka Emika Panda robot

In the following, we will refer to the frame $\Sigma_0$in Figure 1 as base frame of the respective robot. The overall setup is shown in Figure 2 and the world frame is assumed to coincide with the base frame of the left robot. The following transformation matrix $T_w^{0_r}$ apply to pass from the base frame of the right robot to the world frame

$$T_w^{0_r} = \begin{bmatrix} 1 & 0 & 0 & -0.01 \\ 0 & 1 & 0 & -0.83 \\ 0 & 0 & 1 & 0.03 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

where $I_n \in \Re^{n \times n}$ is the $(n \times n)$ identity matrix and $0_{n\times m} \in \Re^{n \times m}$ is the $(n \times m)$ zero matrix. The head is mounted on a podium stand in the middle of the workspace and its position

[1] https://www.franka.de/Panda_Datasheet_May_2019.pdf

(represented by the upper point along the axis of symmetry) is $p_h = [0.57 \quad -0.46 \quad 0.38]^T$ m for starting configuration **[gr2]** and $p_h = [0.49 \quad -0.46 \quad 0.38]^T$ m otherwise.

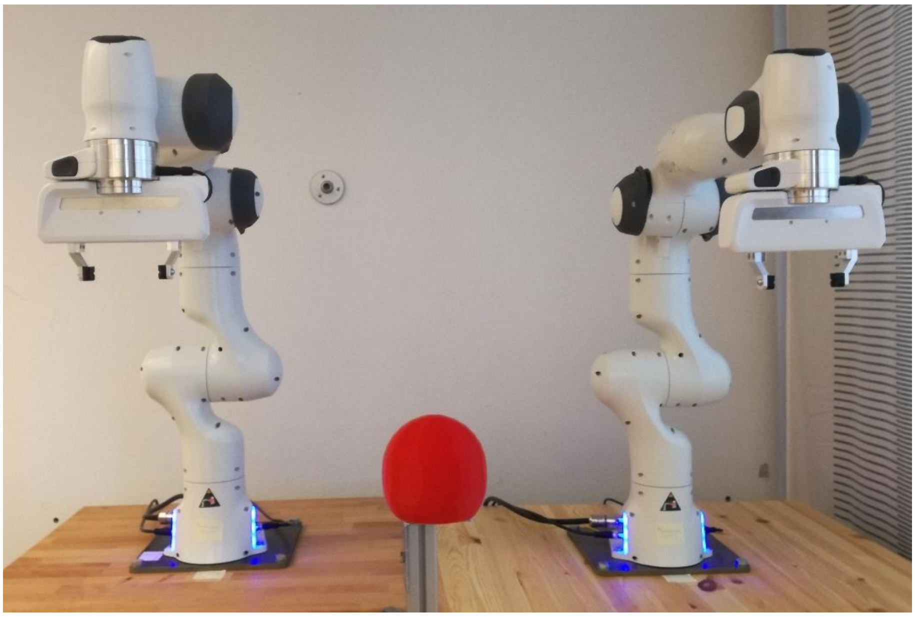

Figure 2 - Dual-arm setup

End effector

Each robot is equipped with parallel grippers, shown in Figure 3, and provided by Franka Emika. Customized long fingers, reported in Figure 4, have been adopted for left robot in case of starting configurations **[pg1]**, **[cr]**, **[ft]** and **[fd].**

Figure 3 - Robot gripper with standard fingers from Franka Emika

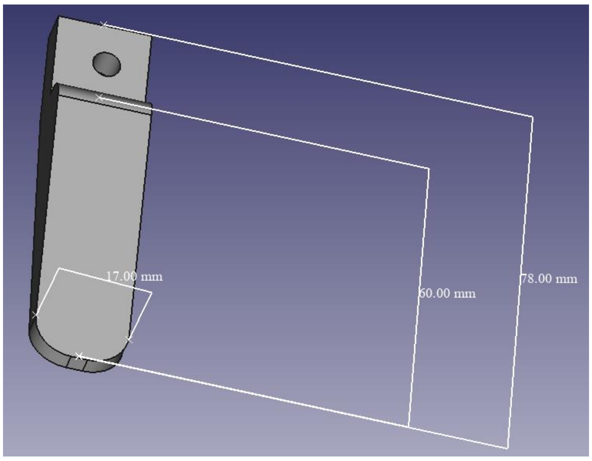

Figure 4 - Customized fingers

Sensors

The following sensors are adopted:

- Logitech C920 PRO HD Webcam[2] positioned over the setup;
- Link-side torque sensors in all 7 axes for each robot.

T-shirt

The T-shirt is reported in Figure 5 and has the following characteristics:

- Producer: Zara
- Type: Woman L
- Color: salmon
- Material: 100% cotton
- Product id: 4424/308/633/L
- Measurements:

A	B	C	D	L_n
19 cm	50 cm	2.9 cm	10 cm	50.9 cm

Figure 5 - Used T-shirt

[2] http://static.highspeedbackbone.net/pdf/Logitech%20C920%20HD%20Pro%20Webcam%20Data%20Sheet.pdf

The proposed approach is also validated with a second T-shirt, shown in Figure 6, with the following characteristics:

- Producer: H&M
- Type: Man S regular fit
- Color: blue
- Material: 100% cotton
- Product id 0685816
- Measurements:

A	B	C	D	L_n
15.4 cm	50 cm	2.5 cm	8 cm	40 cm

Figure 6 - Second T-shirt for validation

Workstation specifications

The employed PC is equipped with Intel(R) Core(TM) i7-8700 CPU@3.2 GHz and 16GB RAM. The operative system is Ubuntu 18.04 (Bionic Beaver) with real-time kernel 4.18.16-rt9. The library libfranka version 0.5 is used for robots control.

Software architecture

A ROS-based software architecture is employed with ROS Melodic. The overall software architecture in terms of ROS nodes and exchanged data is summarized in Figure 7.

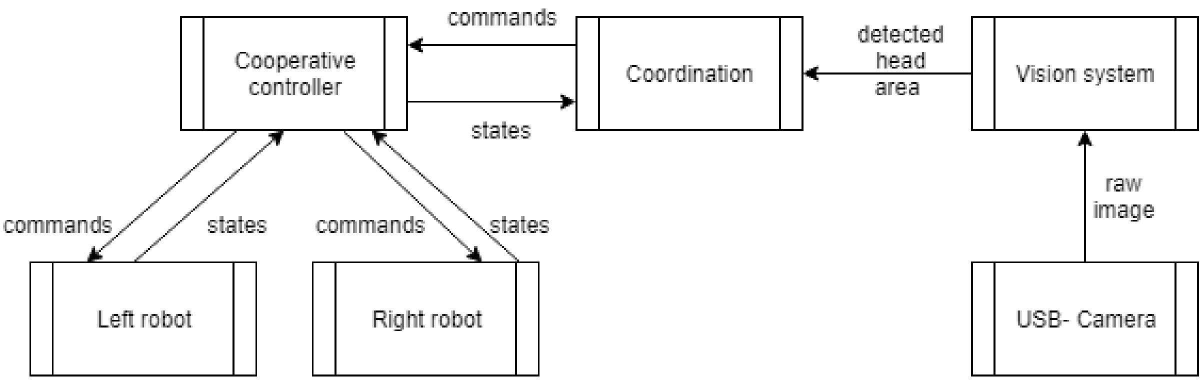

Figure 7 - ROS architecture

Results

What makes the system successful?

The system succeeds by performing a human inspired dressing strategy that is separated into multiple phases: **[GR1]** (grasping the shirt in one point), **[GR2]** (grasping the shirt in a second point while having it grasped in one point already) and **[MAN]** manipulating the T-shirt in order to fulfill the task. The manipulation phase is composed of the following steps:
Step 1: Prepositioning of the T-shirt. The team moves the centroid over the known position of the head while stretching the T-shirt;
Step 2: Detection of head using force measures. The team moves down while monitoring the change in estimated forces in Z direction. This results in the team detecting when there is contact between the garment and the head; in the moment of contact, the team stops the downward motion.
Step 3: Dressing strategy. The vision system splits the circle associated with the head into two halves and measures the free area in each of them. These measures are provided to the team (see Figure 8). The team motion is guided by the difference of those two area measures and tries to minimize the difference. Additionally, the team performs random motion along X, Z directions and around Z axis. The distance of the team is also gradually decreased until a threshold is reached which triggers the stretching of the shirt and the reset of the formation distance. During the entire manipulation, the exerted forces are monitored and, if a given threshold is exceeded, the team resets the shirt to a reset position above the head and restarts the strategy from step 2.The restart strategy is also activated if the elapsed time exceeds a predefined value.

Figure 8 - Output of the vision system

As soon as the free area of both halves of the head exceeds a predefined threshold, a downward motion is initiated. The team moves down until they reach a position below the head which concludes the dressing task.

What makes the system fail?

The system is dependent on the quality of the vision system. If the illumination or the color of head changes the received information is unusable for the team.
The vision also does not consider the neck opening, which means that the team will move down given enough head area is visible even when the head is not fitted through the neck hole of the shirt.
The manipulation can also not handle if the opening of the shirt is not fitted over the head given a grasping that results in a long hanging side of the T-shirt like in starting configurations **[fd]** or **[ft]**.

What was improved compared to other methods?

Cloth manipulation and more precisely dressing as a research field has only recently emerged. Dressing tasks are often customised to the methods employed and it is difficult to transfer approaches between different works as the assumed assumptions can vary a lot.
A work most closely related to the proposed benchmark task is shown in [Reinforcement Learning of Clothing Assistance with a Dual-arm Robot] where a T-shirt is dragged over a mannequin with the arms already positioned through the sleeves of the shirt. The manipulators drag the shirt from the back of the mannequin over the head. A motion capture system is used to detect the position of topological features that are used in the reward function for the proposed reinforcement learning approach.
This approach could be adopted to tackle the proposed dressing task but requires additional motion capture sensors.
The presented baseline implementation is designed in a way to make it easy to reimplement while still highlighting the challenges of the task. The biggest strength of the baseline is also its downfall regarding the more challenging starting configurations: the simplicity of the approach using only a simple usb-camera as guidance for the task.
This simplicity however leads to a limitation in regards to the different T-shirt poses the baseline can deal with as it is evident with the performances for the **[ft]** and **[fd]** starting configurations.

Chosen grasping points and/or grasping strategy

The devised manipulation strategy considers that the T-shirt is grasped at two points along the neck collar and with as much as possible fabric inside the fingers. In this way, the amount of T-shirt below the fingers is minimized.
Depending on the starting configuration, different grasping strategies are carried out:

- **[pg2]**: there is no grasping phase and the user puts the T-shirt as described in the above;
- **[pg1]**: the left robot starts from a predefined configuration where the user manually performs the first grasping on the neck collar of the T-shirt while trying to put one sleeve between the fingers as represented in Figure 9 – left figure. The automatic grasping of the second point is performed by first moving the right robot into a predefined pose such that the T-shirt collar is inside the fingers, then a downward motion is performed until the end of the collar and finally the fingers are closed;
- **[cr]:** the left robot moves to a predefined position close to the table where the T-shirt is crumpled and tries the grasping. Because of the crumpled configuration, this generally results in having much fabric inside the finger that can then facilitate the manipulation strategy. The left robot then moves to the predefined position used as starting configuration for the starting configuration **[pg1]** and the same strategy as in **[pg1]** is adopted;
- **[ft]:** the left robot moves to a predefined position close to the table where the T-shirt is positioned flat and tries the grasping. The T-shirt position on the table is known and the predefined position is defined so that the first grasping point is on the neck collar close to the shoulder seam as reported in Figure 9 – right figure. The left robot then moves to the predefined position used as starting configuration for the starting configuration **[pg1]** and the same strategy as in **[pg1]** is adopted;
- **[fd]:** the same strategy as in **[ft]** is carried out.

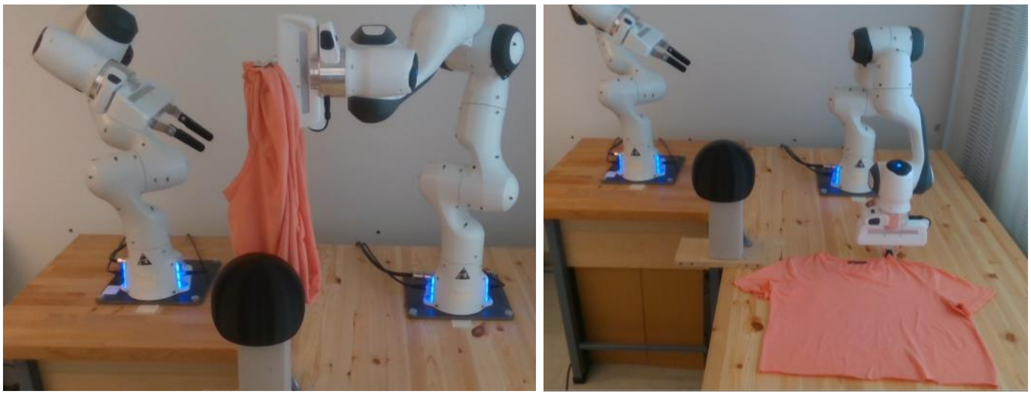

Figure 9 - Grasping poses