

# Door opening and traversal with an industrial cartesian impedance controlled mobile robot

Marvin Stuede, Kathrin Nuelle, Svenja Tappe, Tobias Ortmaier<sup>1</sup>

**Abstract**—This paper presents a holistic approach for door opening with a cartesian impedance controlled mobile robot, a KUKA KMR iiwa. Based on a given map of the environment, the robot autonomously detects the door handle, opens doors and traverses doorways without knowledge of a door model or the door's geometry. The door handle detection uses a convolutional neural network (CNN)-based architecture to obtain the handle's bounding box in an RGB image that works robustly for various handle shapes and colors. We achieve a detection rate of 100% for an evaluation set of 38 different door handles, by always selecting for highest confidence score. Registered depth data segmentation defines the door plane to construct a handle coordinate frame. We introduce a control structure based on the task frame formalism that uses the handle frame for reference in an outer loop for the manipulator's impedance controller. It runs in soft real-time on an external computer with approximately 20 Hz since access to inner controller loops is not available for the KMR iiwa. With the approach proposed in this paper, the robot successfully opened and traversed for 22 out of 25 trials at five different doors.

## I. INTRODUCTION

Mobile service robots need the ability to open doors in order to gain full autonomy. Therefore, opening doors is a widely addressed topic in research, consisting of two subproblems: (a) locating the door and handle and (b) the opening process itself.

Detection usually works either range based or vision based. Recent works e.g. use 3D colored laser scans [1] or deep learning based methods [2]. Robust door detection remains a challenging task due to the variety in color and contrast between the closed door and the wall. Detecting the handle possesses similar difficulties and often additionally includes problems with incorrect depth data caused by reflecting surfaces. A classical approach for vision based detection in RGB images is canny edge detection of a handle image as in [3][4]. For this approach a high contrast of the handle to the surrounding is necessary, resulting from its geometric shape and color. For low contrast problems, a common approach are Haar feature based cascade classifiers (HCCs) that lead to detection rates of over 90% [5][6]. To discard false positive results, post-processing of HCC is usually executed e.g. by K-means clustering [6] or template matching [7]. A hybrid approach supplements image-based detection by 3D information as done by Meeussen *et al.* [8]. Their laser-based 3D perception in combination with

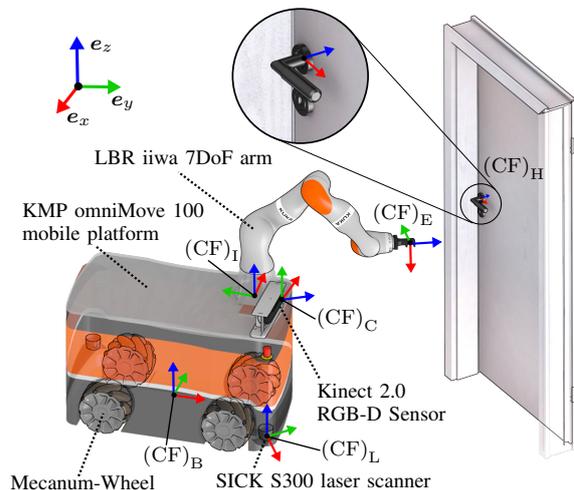


Fig. 1: The components of the *KUKA KMR iiwa* robot, which was used for door opening, and the poses of the corresponding coordinate frames.

HCC based 2D perception achieves a 60% detection rate for arbitrary door handles.

Recently, deep learning based methods are used, as in [2]. They use a CNN together with RGB image and RANSAC-based point cloud segmentation to determine the depth information of the handle and achieve a 97% detection rate. However, a correct detection depends on successful segmentation by K-means clustering, so it remains unclear how robust this approach executes for similar handle and door color.

Several approaches for the opening of doors were proposed since the early 1990s [9]. Early approaches made assumptions about the doors size and hinge position. Force feedback with active [10] or passive [11] compliance enables door opening in a reactive manner for a more general approach. To define a specific manipulator movement, an online estimation or sensing of geometric parameters such as the door radius or center of rotation proved useful. Following strategies for estimation exist: based on the prior movement of the end-effector and corresponding path regression [12][13] or force and velocity measurements [14], respectively. One state of the art approach presented in [15] utilizes a 6-axis force-torque sensor at the end-effector to estimate the motion direction of different doors and drawers. Especially for motion estimating approaches, relative motion between the robot and handle must be avoided. Therefore, a gripper is required as the robot end-effector.

\*This work was supported by Roboterfabrik

<sup>1</sup>All authors are with the Institute of Mechatronic Systems, Leibniz Universität Hannover, D-30167 Hannover, Germany [marvin.stuede@imes.uni-hannover.de](mailto:marvin.stuede@imes.uni-hannover.de)

It should be noted that the presented approaches require control frequencies greater than 100 Hz and, therefore, low level system access [10][13][14][15], which is not inherently available for commercial robots. A control frequency as low as 10 Hz was achieved in [16] by employing equilibrium point control and [17] as part of the TREX control framework. However, in the first approach only pulling drawers and cabinet doors to open is investigated and no statement about total opening time is made. The second approach focuses more on the framework than development of specific door opening controllers. Control strategies for commercial robots without low level access still need further research.

This paper presents an approach for robust door handle detection in varying conditions as well as a control structure that can be used as an outer loop for a commercial cartesian impedance controlled mobile robot. The remainder of this paper is organized as follows: section II begins with a problem statement and description of the mobile system. The opening process is furthermore described by means of navigation to the door, opening state and handle detection and the control structure for opening. In section III the proposed methods for door handle detection and door opening are evaluated and section IV summarizes the findings.

## II. DOOR OPENING PROCESS

### A. Problem statement

We investigate the problem of an autonomous robot, which is operating in an indoor environment consisting of rooms separated by doors. The aim is to send arbitrary goal poses within the unmodified environment, so that the robot opens every closed door on the path to the goal. We subdivide the door opening problem into door handle detection and door opening, not including door localization, and formulate the following requirements for the solution:

- A map of the environment is priorly obtained as a 2D occupancy grid map, eg. by grid based SLAM [18].
- The robot must detect any kind of rotating lever latch handle in order to successfully grasp the lever, regardless of handle and door color or shape.
- For door opening, we restrict to the consideration of non-transparent doors, openable by pushing.

Door detection is not considered in this work, hence the door positions in the map are given by two points  ${}^{(M)}\mathbf{h}_i, {}^{(M)}\mathbf{f}_i \in \mathbb{R}^2$ , both defined in the map coordinate frame  $(CF)_M$ , to represent the hinge and stop positions and a value  $w_i \in \{0, 1\}$  to denote the opening direction (see Fig. 2a).

### B. Prerequisites of the mobile system

As the mobile system, we assume a redundant torque controlled arm mounted on a holonomic base (such as *KUKA KMR iiwa*, see Fig. 1). Attached to the arm is a simple hook with a parabolic opening to avoid longitudinal movement of the handle during door opening. For range based localization and navigation, two *SICK S300* laser scanners are used. Additionally necessary is a base-mounted RGB-D Camera (e.g. a *Microsoft Kinect 2.0*), facing the front. On application level, the arm can either be position- or impedance controlled, with

the latter especially allowing for cartesian impedance control in the form

$$\mathbf{f}(t) = \mathbf{D}\Delta\dot{\mathbf{x}}_E + \mathbf{K}\Delta\mathbf{x}_E, \quad (1)$$

with stiffness matrix

$$\mathbf{K} = \text{diag}(k_x, k_y, k_z, k_\phi, k_\psi, k_\theta) \quad (2)$$

and damping matrix

$$\mathbf{D} = \text{diag}(d_x, d_y, d_z, d_\phi, d_\psi, d_\theta). \quad (3)$$

The end-effector pose  $\mathbf{x}_E = (x_E, y_E, z_E, \phi_E, \psi_E, \theta_E)^T$  is defined with RPY convention. We exclusively use the arm in cartesian impedance control mode, only modifying  $\mathbf{K}$  for the individual steps of the opening process, and a joint velocity  $\dot{\mathbf{q}}_{\text{ref}}$  as the input for the robot's control system. The described methods are implemented on an external notebook computer, running *ROS kinetic* and commanding to the robot controller via Ethernet in a frequency of 20 Hz.

### C. Navigation to the door

Before passing a doorway, it must be determined whether the door is opened or closed. Therefore, the mean cost value  $\bar{c} = \frac{1}{n} \sum_{c_k \in \mathcal{A}} c_k$  of costs  $c_0 \dots c_n$  in the local costmap is calculated in an area  $\mathcal{A}$  within the door frame (see Fig. 2a). The mean cost of closed doors is significantly higher as of open doors. Hence, the opening state is determined by thresholding (e.g.  $\bar{c} \geq 90$ ) with an additional search for a door handle to avoid detecting the door as closed if obstacles are within the doorway.

Opening a closed door requires the robot's front to be accurately aligned to the door, especially when the robot is only slightly narrower than most doorways, which applies to the *KMR iiwa*. To compensate uncertainties of global localization using AMCL [19], laser scan measurements in an area in front of the robot correct its position. A measurement subset  $\mathcal{L}_s$  is used for a RANSAC based line segmentation to obtain a line  $\mathcal{G}$  which corresponds to the door's front. Based on this line, a distance  $l$  and angle  $\gamma$  are calculated (see Fig. 2b) and defined in the robot's base frame  $(CF)_B$ . These variables are then used as control variables to position the robot at a defined distance  $l_d = 0.7$  m and angle of  $\gamma_d = 0$  to the door.

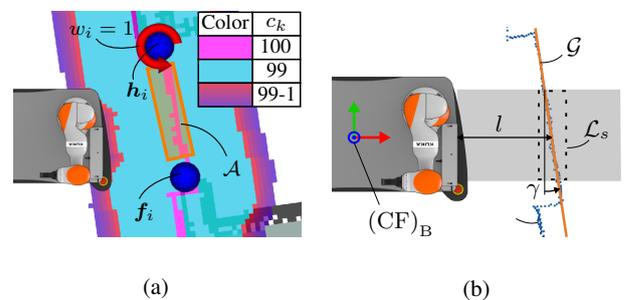


Fig. 2: Aligning to a closed door with the determination of the door opening state in (a) and the base positioning procedure in (b).

#### D. Detecting and Grasping the handle

The main challenge in door handle detection based on range measurements is erroneous depth estimation because of a specular handle surface. To overcome this problem, the handle is detected in the RGB image and tactile feedback of the arm is employed to determine its distance to the door. The detection in the RGB image results in a bounding box around the handle. For this, we use a CNN based architecture proposed by Huang *et al.* [20]. The network consists of a Resnet 101 feature extractor to create box proposals together with a Faster R-CNN network for object detection. Based on the available model checkpoint, we train the network with 200 manually photographed images of differently shaped door handles with variation in camera angle and illumination. For an evaluation set of 60 images of 38 different door handles we achieve 0.95 mAP@0.5IoU; an excerpt of the set is depicted in Fig. 3. In the rare case of detecting several objects in one image as door handle, we primarily select the box with the highest confidence score and secondary based on largest area. This strategy leads to a detection rate of 100% on the evaluation set and coincides with our experience during application in practice.

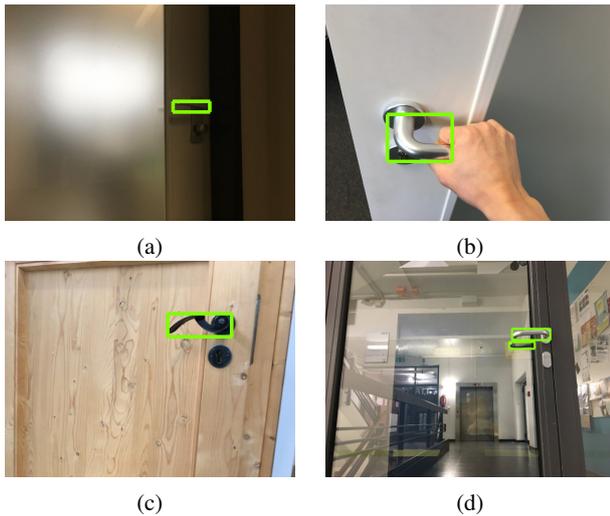


Fig. 3: Detected door handles for different scenarios. Green boxes indicate the output of the object detection network. Training was based on similar pictures as these examples.

In the bounding box, we determine the intersection between the handle's rotational axis and the door plane in the image, using the approach proposed by Klingbeil *et al.* [6]. For a bounding box given by the top left corner  $(x, y)$ , height  $h$  and width  $w$ , the intersection point is approximately located at  $(x + \alpha w, y + \beta h)$ , where  $\alpha$  and  $\beta$  are empirically chosen constants, assumed to be valid for every handle.

The registered point cloud of the Kinect sensor is used to obtain the door plane. A rectangular mask  $M(x, y, h, w)$  half as wide and double as high as the bounding box of the handle located directly above the handle's bounding box is applied to the point cloud and extracts the points used for RANSAC-based plane segmentation. We choose this position

because the area directly above the handle is usually planar and non-transparent. When a small distance threshold for RANSAC inliers is chosen (e.g. 3 mm), the segmentation is also applicable to doors containing glass panels. With regard to handle detection, this method therefore offers an advantage over 2D-laser scan based detection. The intersection between door plane and handle rotational axis defines the origin for a handle coordinate frame  $(CF)_H$ . The frame's  $z$ -axis  ${}_{(H)}e_z$  points normal to the door plane and the  $y$ -axis  ${}_{(H)}e_y$  perpendicular to the ground.

To grasp the handle, the end-effector is moved to a centered position  ${}_{(H)}(70 \text{ mm}, 0, -100 \text{ mm})^T$  in front of the handle (see Fig. 4a). The 70 mm originate from a German standard regulating the length of door handles to an interval from 93 mm to 140 mm [21]. We then move the end-effector forward until a cartesian force threshold  ${}_{(E)}f_z \geq 5 \text{ N}$  is exceeded. A consecutive upward and forward motion then places the opening of the grasping hook above the handle. A downward motion with low cartesian stiffness in  ${}_{(E)}z$  direction lets the tool then slide on the handle, where the stopping condition is again defined as a cartesian force threshold  ${}_{(E)}f_x \geq 10 \text{ N}$ . Since the rotational axis of the handle is given by  ${}_{(H)}e_z$ , the handle is turned by a circular motion around this axis (see Fig. 4b). Potential uncertainties in the rotational axis pose are compensated by choosing a low rotational stiffness  $k_\theta$ , which will always keep the handle in contact. By reaching an empirically chosen maximum force  ${}_{(E)}f_x = 35 \text{ N}$  the circular motion stops.

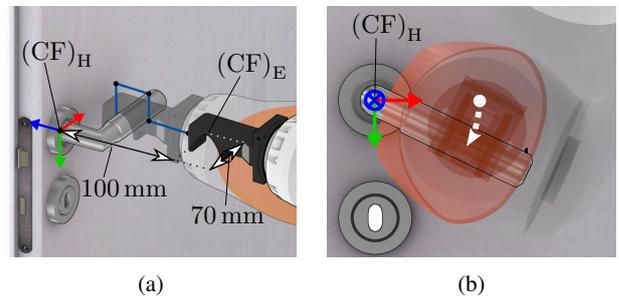


Fig. 4: Grasping and turning the handle with tactile feedback. The blue path in (a) indicates the path of the tool center point, which is given as  $(CF)_E$ . The white path in (b) indicates the circular path to turn the handle.

#### E. Opening the door

The process of opening a door can be described as operating a compliant mechanism with one degree of freedom, given by the door's rotational joint. To manipulate this kind of mechanism, a common approach for controller design is the task frame formalism, proposed by Mason [22], extended by Bruynickx and De Schutter [23] and implemented for the door opening problem in [8][24]. Opening the door is described in terms of a compliant motion, velocity and force references in a task frame. Therefore, the task frame is chosen to be the handle frame  $(CF)_H$  and the task is defined by a constant velocity/force reference. The selection makes

the use of a gripper for force based direction estimation as in [24] unnecessary because the rotation of  $(CF)_H$  around  ${}^{(H)}e_y$  is continuously tracked by laser scan RANSAC line segmentation, resulting in a homogeneous transformation  ${}^L T_H$ . Tracking the frames position is not needed due to the control structures design; choosing the handle frame as the task frame therefore mainly serves illustrative purposes.

We phrase the objective of the control structure as *Open the door using the handle with a defined, constant velocity* and propose a structure based on Prats Framework [25] consisting of three separate controllers:

1. A feed forward task controller to open the door with constant velocity.

2. A grasp controller to keep the robot in contact with the handle.

3. A base controller to move the platform through the doorway and simultaneously avoid collision with the door.

By this modular approach, the partial solutions are separated to simplify adapting the controllers to different robotic systems.

1) *Task controller*: Tracking  $(CF)_H$  ensures that the task can be described as a velocity reference

$${}^{(H)}\dot{\mathbf{x}}_{t,\text{ref}} = (0, 0, \dot{z}_{\text{ref}}, 0, 0, 0)^T \quad (4)$$

defined in  $(CF)_H$  with  $\dot{z}_{\text{ref}}$  being a constant scalar velocity. The end-effector velocity is then given by

$${}^{(E)}\dot{\mathbf{x}}_t = {}^E \mathbf{W}_H ({}^{(H)}\dot{\mathbf{x}}_{t,\text{ref}} - {}^H \mathbf{W}_{B(B)} \dot{\mathbf{x}}_B), \quad (5)$$

where  ${}^E \mathbf{W}_H$  and  ${}^H \mathbf{W}_B$  are twist transformation matrices.  ${}^E \mathbf{W}_H$  is associated to the homogeneous transformation  ${}^E T_H$  from  $(CF)_H$  to  $(CF)_E$ , in our case defined as

$${}^E T_H = {}^E T_L {}^L T_H \quad (6)$$

with a transformation between laser frame and end-effector frame  ${}^E T_L$ , known by the arm's forward kinematics and the construction data of the robot.  ${}^L T_H$  is estimated online by the laser scan based approach previously described.  ${}^H T_B$  follows accordingly as a transformation between the robot's base and the handle frame. Subtracting the bases current velocity  $\dot{\mathbf{x}}_B$  in Fig. 5 ensures that only the task velocity reference is transferred to the door handle.

2) *Grasp controller*: A safe and continuous door opening requires the grasping tool to maintain contact with the handle during the whole process. This means that the pose  ${}^{(H)}\mathbf{x}_E$  of the end-effector frame  $(CF)_E$  should remain constant with respect to the handle frame  $(CF)_H$ . A control deviation for the rotational degrees of freedom can therefore be defined as a difference between the pose  ${}^{(H)}\mathbf{x}_{E,\text{ref}}$  after finishing handle turning and the current pose  ${}^{(H)}\mathbf{x}_E$ :

$${}^{(H)}\mathbf{e}_r = {}^{(H)}\mathbf{x}_{E,\text{ref}} - {}^{(H)}\mathbf{x}_E. \quad (7)$$

Since the position of the handle frame  $(CF)_H$  on the door plane is unknown during door opening, the translational degrees of freedom of  $\mathbf{x}_E$  are not controlled with a position controller. Instead, we use a force controller to ensure that the handle is continuously pressed down. The control error

$${}^{(E)}\mathbf{e}_f = {}^{(E)}\mathbf{f}_{x,\text{ref}} - {}^{(E)}\mathbf{f}_x \quad (8)$$

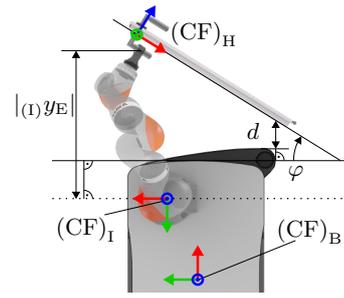


Fig. 5: Geometric relation of the door opening process:  $|{}_{(1)}y_E|$  is the distance between the end-effector frame  $(CF)_E$  and manipulator base frame  $(CF)_I$ . The door opening angle  $\varphi$  is defined as the angle between the  $y$ -axis of the base frame  $(CF)_B$  and  $x$ -axis of  $(CF)_H$ .

follows from the cartesian force  ${}^{(E)}\mathbf{f}_x$  and a reference force  ${}^{(E)}\mathbf{f}_{x,\text{ref}}$ . The grasp controller therefore is implemented as a hybrid structure with a rotational controller  $\mathbf{c}_r(\mathbf{e}_r) \in \mathbb{R}^6$  and a force controller  $\mathbf{c}_f(\mathbf{e}_f) \in \mathbb{R}^6$ , resulting in an end-effector velocity:

$${}^{(E)}\dot{\mathbf{x}}_{\text{gr}} = \mathbf{S}_r {}^E \mathbf{W}_H \mathbf{c}_r(\mathbf{e}_r) + \mathbf{S}_f \mathbf{c}_f(\mathbf{e}_f). \quad (9)$$

Note that by choosing the selection matrices

$$\mathbf{S}_r = \text{diag}(0, 0, 0, 1, 1, 1), \quad \mathbf{S}_f = \text{diag}(1, 0, 0, 0, 0, 0) \quad (10)$$

the  $y$ - and  $z$ -direction in  $(CF)_E$  are ignored. The movement along the handle in  $y$ -direction is inherently avoided by setting a small cartesian stiffness  $k_y$ , which keeps the tool in place by solely controlling in  $x$ -direction. The  $z$ -direction must not be controlled here because it corresponds to the task direction.

Both the task- and grasp controller result in a velocity reference, which is combined to a total velocity and transformed to the manipulator base frame  $(CF)_I$ :

$${}_{(1)}\dot{\mathbf{x}}_{\text{tot}} = {}^I \mathbf{W}_E ({}^{(E)}\dot{\mathbf{x}}_t + {}^{(E)}\dot{\mathbf{x}}_{\text{gr}}). \quad (11)$$

This velocity is then transformed with the Jacobian's Moore-Penrose inverse  $\mathbf{J}^+(\mathbf{q})$  to obtain the joint velocity

$$\dot{\mathbf{q}}_{\text{ref}} = \mathbf{J}^+(\mathbf{q}) {}_{(1)}\dot{\mathbf{x}}_{\text{tot}} + (\mathbf{I} - \mathbf{J}^+(\mathbf{q})\mathbf{J}(\mathbf{q})) \frac{\partial h}{\partial \mathbf{q}}, \quad (12)$$

where  $(\mathbf{I} - \mathbf{J}^+(\mathbf{q})\mathbf{J}(\mathbf{q}))$  is a nullspace projector to execute a secondary task in the manipulator's nullspace using the gradient projection method [26]. In this case, a cost function  $h(\mathbf{q})$  is optimized to avoid mechanical joint limits [27].

3) *Base controller*: The spatial constraint of the manipulator's workspace requires the mobile base to continuously move through the doorway during the door opening process. We only utilize one degree of freedom of the mobile base, the  $x$ -direction of the base frame  $(CF)_B$  (see Fig. 5), although a secondary task such as manipulability increase could be projected in the remaining degrees of freedom. The base is moved through the doorway by a controller that uses the deviation

$$\mathbf{e}_b = {}_{(1)}y_{E,\text{ref}}(\varphi) - |{}_{(1)}y_E| \quad (13)$$

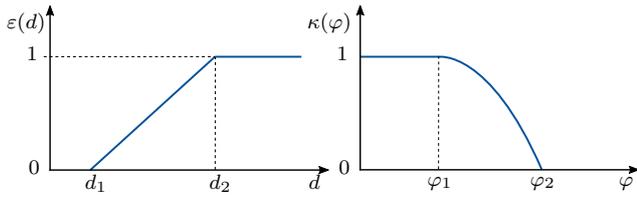


Fig. 6: Factors  $\varepsilon$  and  $\kappa$  to avoid collisions with the door.  $d_{1,2}$  and  $\varphi_{1,2}$  must be chosen depending on the robot's geometry.

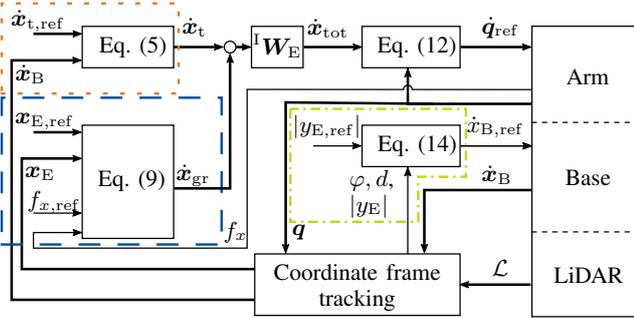


Fig. 7: The control structure, consisting of the **task controller** (dotted), **grasp controller** (dashed) and **base controller** (dash-dotted).

with  $|{}_{(I)}y_E|$  being the current distance between manipulator base and end-effector link and  ${}_{(I)}y_{E,ref}(\varphi)$  as a reference, which is a piecewise-defined linear decreasing function with  ${}_{(I)}y_{E,ref}(0) = |{}_{(I)}y_E|$ . Decreasing  ${}_{(I)}y_{E,ref}$  for an increasing door angle  $\varphi$  is necessary due to the arm's limited workspace. The base velocity is then calculated as

$${}_{(B)}\dot{x}_{B,ref} = \kappa(\varphi)\varepsilon(d)c_b(e_b) \quad (14)$$

where  $c_b(e_b)$  is a position controller. To avoid collisions with the door, two additional factors  $\varepsilon(d)$  and  $\kappa(\varphi)$  are introduced which depend on the critical distance to the door  $d$  and the current door opening angle  $\varphi$  (see Fig. 5). The functions are designed so that  $\lim_{d \rightarrow 0} \varepsilon = 0$  and  $\lim_{\varphi \rightarrow 90^\circ} \kappa = 0$  ensure that the robot stops close to the door and when the door is opened. Both functions also hold  $\max(\varepsilon, \kappa) \leq 1$  to only decrease the velocity (see Fig. 6). Fig. 7 depicts the complete control structure. This structure is implemented as a node in *ROS* and set to a cycling frequency of 20 Hz.

### III. EXPERIMENTAL RESULTS

The independent solutions to the sub-problems door handle detection and door opening are evaluated separately. Additionally, we test the overall door opening process for several different doors.

#### A. Handle detection

To evaluate the door handle detection, we position the robot three times each in front of twelve different unknown<sup>1</sup> doors, perform handle frame detection, move the end-effector

<sup>1</sup>Unknown, as that the door handles were not used for training the handle detector

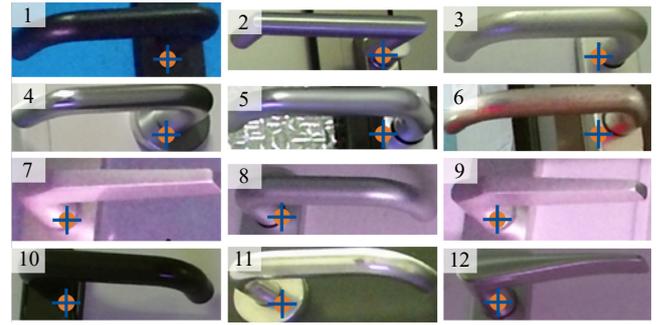


Fig. 8: Estimated intersection (blue cross) between handle axis and door plane for twelve different unknown doors. The images are cropped in the size of the bounding boxes resulting from the neural network.

Direction	Mean absolute error
x	8.6 mm $\pm$ 1 mm
y	12.1 mm $\pm$ 1 mm
z	6.5 mm $\pm$ 1 mm

TABLE I: Mean absolute error of the handle frame estimation for 33 measurements at 11 different doors. The  $xyz$ -directions correspond to the base vectors of  $(CF)_H$ .

to a specific pose defined in the handle frame and manually measure the actual distance to the handle origin. Handle axis detection in the RGB images results in the estimated poses depicted in Fig. 8. For eleven of the twelve doors, our method finds a reasonable handle frame, resulting in the mean absolute error listed in Tab. I. The plane segmentation fails only for one door (5 in Fig. 8) containing a fluted milk glass window, which results in a depth measurement error of several centimeters. Otherwise, our method delivers robust results, even for dim environments (12), low color contrast (3) or glass doors (6). An error in the magnitude of the values in table I can be well compensated in  $y$ - and  $z$ -direction by the tactile feedback when grasping the handle. Therefore, the handle can be grasped successfully in 33 out of 36 trials.

#### B. Door opening

The main objective of the door opening controller is to open the door with a defined, constant velocity, which must therefore result in a linear increase of the door opening angle  $\varphi$ . We evaluate this by performing automated door opening on five different doors (five times each), from 0.81 m to 0.98 m wide, by measuring the door opening angle  $\varphi$  with the described laser scan based method. The controllers  $c_b$  and  $c_t$  are implemented as P- and  $c_f$  as a PD controller. The evaluated doors have spring loaded door handles, therefore the D-Term in  $c_f$  is used to decrease oscillations during compliant contact. Fig. 9 depicts  $\varphi$  for a reference velocity of  $\dot{z}_H = 50$  mm/s. In every measurement,  $\varphi$  shows a clear linear increase with the door's width mainly influencing total opening time. As expected from the feed forward task controller (Eq. (5)), the mean opening velocity  $\bar{z}_H$  stays below  $\dot{z}_H$ . Increasing the reference velocity up to  $\dot{z}_H = 90$  mm/s still leads to opening with constant velocity (see Fig. 10),

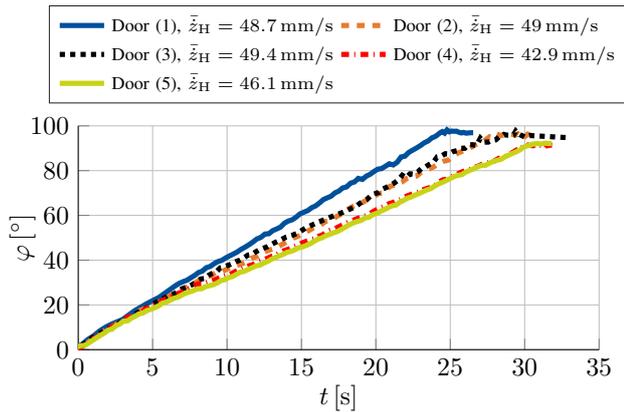


Fig. 9: Door opening angle  $\varphi$  over time  $t$  for five different doors with mean opening velocity  $\bar{z}_H$ . Door (1) is 0.81 m wide, opens clockwise (cw), Door (2) is 0.93 m wide, opens cw, Door (3) is 0.97 m wide, opens cw, Door (4) is 0.93 m wide, opens counter clockwise (ccw), Door (5) is 0.98 m wide, opens ccw.

although also increasing oscillations of  $f_x$  (up to  $f_{x,\max} = 70$  N compared to  $f_{x,\max} = 38$  N for  $\dot{z}_H = 50$  mm/s) which could ultimately lead to a failed opening attempt. Therefore, we choose a reference velocity of  $\dot{z}_H = 50$  mm/s achieving 22/25 successful opening attempts. An attempt is considered successful if the robot opens the door to at least 90 degrees without losing contact to the handle. The success rate of 88% is comparable to state-of-the-art approaches as in [6] (91%), [4] (60%) or [28] (52%). Although it must be noted that all approaches apply different evaluation settings as the number of different doors, door size or only consider opening without traversal.

### C. Overall process

Our introduced door opening process consists of the subsolutions for opening state detection, handle detection and door opening. It is implemented as an extension to the *ROS navigation stack*. Instead of sending goals to the navigation stack directly, the proposed system checks whether the global path to a certain goal passes through doorways, commands navigation before any existing doorways and, if necessary, triggers the opening procedure. An exemplary process of driving into a room with a closed door is given in Fig. 11 with a temporal evaluation for each consecutive step ( $\dot{z}_H = 70$  mm/s). The robot aligned itself to the door after 10s, followed by door handle detection which averagely takes 0.3s and handle grasping within 20s. Door opening of this door is finished within 24s.

Consequently, the whole process starts with aligning and ends after passing the doorway and is completed after 75s (81s for  $\dot{z}_H = 50$  mm/s). We did not optimize the system for speed, therefore the overall duration could further be decreased, especially by adjusting the aligning threshold and maximum cartesian velocity of the arm, which is currently limited to 100 mm/s.

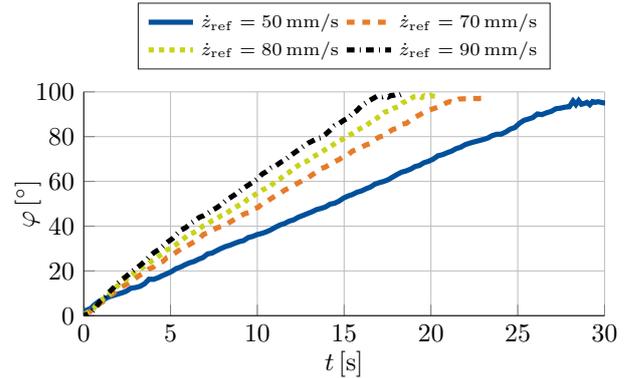


Fig. 10: Approximately linear increasing door opening angle  $\varphi$  over time  $t$  for door (2) using different velocities  $\dot{z}_{ref}$ . Oscillations increase for higher velocities.

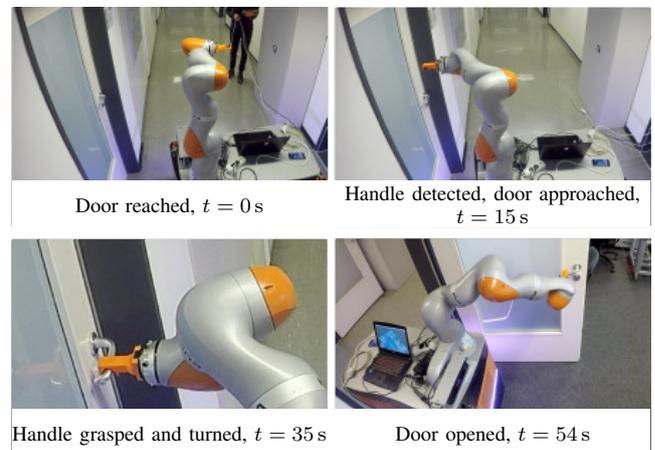


Fig. 11: Driving into a room with a closed door. Time evaluation starts with checking whether the door is closed and ends after the doorway is passed.

## IV. CONCLUSION

The goal of this paper was to demonstrate an approach to reliably open and traverse doors with an industrial mobile robot that doesn't allow for access to inner control loops. The approach was tested in an unaltered office environment, without presumptions about the door's kinematics or geometry for opening. Independent evaluations for door handle detection and the opening algorithm proved that both solutions work robustly: The CNN for handle detection achieves 0.95 mAP@0.5IoU for an evaluation set of 60 images of 38 different door handles and 100% detection rate by selecting based on confidence score. In 36 trials at 12 different doors, the handle was grasped successfully 33 times. The developed control structure for door opening acts as an outer loop to the impedance controller and runs with modest hardware requirements with a frequency of 20 Hz, achieving 22/25 successful openings and traversals at five different doors.

#### REFERENCES

- [1] B. Quintana, S. A. Prieto, A. Adan, and F. Bosche, "Door detection in 3D colored laser scans for autonomous indoor navigation," in *2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. IEEE, oct 2016, pp. 1–8.
- [2] A. Llopart, O. Ravn, and N. A. Andersen, "Door and cabinet recognition using Convolutional Neural Nets and real-time method fusion for handle detection and grasping," in *2017 3rd International Conference on Control, Automation and Robotics (ICCAR)*. IEEE, apr 2017, pp. 144–149.
- [3] S. Kim, H. Cheong, D. H. Kim, and S.-K. Park, "Context-based object recognition for door detection," in *2011 15th International Conference on Advanced Robotics (ICAR)*. IEEE, jun 2011, pp. 155–160.
- [4] B. Axelrod and W. H. Huang, "Autonomous door opening and traversal," in *2015 IEEE International Conference on Technologies for Practical Robot Applications (TePRA)*. IEEE, may 2015, pp. 1–6.
- [5] B. J. Sapp and J. J. Zhang, "Door Handle Detection for the Stanford AI Robot (STAIR)," Tech. Rep., 2005.
- [6] E. Klingbeil, A. Saxena, and A. Y. Ng, "Learning to open new doors," in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, oct 2010, pp. 2751–2757.
- [7] A. Andreopoulos and J. K. Tsotsos, "A Framework for Door Localization and Door Opening Using a Computer Controlled Wheelchair for People Living with Mobility Impairments," in *RSS 2007 Manipulation Workshop*, 2007.
- [8] W. Meeussen, M. Wise, S. Glaser, S. Chitta, C. McGann, P. Mihelich, E. Marder-Eppstein, M. Muja, V. Eruhimov, T. Foote, J. Hsu, R. B. Rusu, B. Marthi, G. Bradski, K. Konolige, B. Gerkey, and E. Berger, "Autonomous door opening and plugging in with a personal robot," *Proceedings - IEEE International Conference on Robotics and Automation*, pp. 729–736, 2010.
- [9] K. Nagatani and S. Yuta, "Designing a behavior to open a door and to pass through a door-way using a mobile robot equipped with a manipulator," in *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'94)*, vol. 2. IEEE, 1994, pp. 847–853.
- [10] C. Ott, B. Bäuml, C. Borst, and G. Hirzinger, "Employing Cartesian Impedance Control for the Opening of a Door: A Case Study in Mobile Manipulation," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2005.
- [11] C. C. Kessens, J. B. Rice, D. C. Smith, S. J. Biggs, and R. Garcia, "Utilizing compliance to manipulate doors with unmodeled constraints," in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, oct 2010, pp. 483–489.
- [12] L. Petersson, D. Austin, and D. Kragic, "High-level control of a mobile manipulator for door opening," in *Proceedings. 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2000)*, vol. 3. IEEE, 2000, pp. 2333–2338.
- [13] T. Winiarski and K. Banachowicz, "Opening a door with a redundant impedance controlled robot," in *9th International Workshop on Robot Motion and Control*. IEEE, jul 2013, pp. 221–226.
- [14] E. Lutscher, M. Lawitzky, G. Cheng, and S. Hirche, "A control strategy for operating unknown constrained mechanisms," in *2010 IEEE International Conference on Robotics and Automation*. IEEE, may 2010, pp. 819–824.
- [15] Y. Karayiannidis, C. Smith, F. E. V. Barrientos, P. Ogren, and D. Kragic, "An Adaptive Control Approach for Opening Doors and Drawers Under Uncertainties," *IEEE Transactions on Robotics*, vol. 32, no. 1, pp. 161–175, feb 2016.
- [16] A. Jain and C. C. Kemp, "Pulling open doors and drawers: Coordinating an omni-directional base and a compliant arm with Equilibrium Point control," in *2010 IEEE International Conference on Robotics and Automation*. IEEE, may 2010, pp. 1807–1814.
- [17] C. McGann, E. Berger, J. Bohren, S. Chitta, B. Gerkey, S. Glaser, B. Marthi, W. Meeussen, T. Pratkanis, E. Marder-Eppstein *et al.*, "Model-based, hierarchical control of a mobile manipulation platform," in *ICAPS Workshop on Planning and Plan Execution for Real-World Systems, Thessaloniki, Greece*, 2009.
- [18] G. Grisetti, C. Stachniss, and W. Burgard, "Improved Techniques for Grid Mapping With Rao-Blackwellized Particle Filters," *IEEE Transactions on Robotics*, vol. 23, no. 1, pp. 34–46, feb 2007.
- [19] D. Fox, "Adapting the Sample Size in Particle Filters Through KLD-Sampling," *International Journal of Robotics Research*, vol. 22, pp. 985–1003, 2003.
- [20] J. Huang, V. Rathod, C. Sun, M. Zhu, A. Korattikara, A. Fathi, I. Fischer, Z. Wojna, Y. Song, S. Guadarrama, K. Murphy, and G. Research, "Speed/accuracy trade-offs for modern convolutional object detectors," in *Computer Vision and Pattern Recognition*, 2017.
- [21] DIN 18255:2002-05, "Türdrücker, Türschilder und Türrossetten - Begriffe, Maße, Anforderungen, Kennzeichnung," 2002.
- [22] M. T. Mason, "Compliance and Force Control for Computer Controlled Manipulators," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 11, no. 6, pp. 418–432, 1981.
- [23] H. Bruyninckx and J. De Schutter, "Specification of Force-Controlled Actions in the "Task Frame Formalism" - A Synthesis," *IEEE Transactions on Robotics and Automation*, vol. 12, no. 4, pp. 581–589, 1996.
- [24] M. Prats, S. Wieland, T. Asfour, A. del Pobil, and R. Dillmann, "Compliant interaction in household environments by the Armar-III humanoid robot," in *Humanoids 2008 - 8th IEEE-RAS International Conference on Humanoid Robots*. IEEE, dec 2008, pp. 475–480.
- [25] M. Prats, A. P. del Pobil, and P. J. Sanz, *Robot Physical Interaction through the combination of Vision, Tactile and Force Feedback*. Springer, Berlin, Heidelberg, 2013.
- [26] B. Siciliano, "Kinematic Control of Redundant Robot Manipulators: A Tutorial," *Journal of Intelligent and Robotic Systems*, vol. 3, pp. 201–212, 1990.
- [27] A. Liegeois, "Automatic supervisory control of the configuration and behavior of multibody mechanisms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 7, no. 12, pp. 868–871, dec 1977.
- [28] T. Ruhr, J. Sturm, D. Pangercic, M. Beetz, and D. Cremers, "A generalized framework for opening doors and drawers in kitchen environments," in *2012 IEEE International Conference on Robotics and Automation*. IEEE, may 2012, pp. 3852–3858.