Accepted Version: 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Oct.23-27, 2022 - Kyoto, Japan.

Pedestrian-Robot Interactions on Autonomous Crowd Navigation: Reactive Control Methods and Evaluation Metrics

Diego Paez-Granados[1†], Yujie He[2], David Gonon[2], Dan Jia[3], Bastian Leibe[3], Kenji Suzuki[4], Aude Billard[2]

***Abstract*— Autonomous navigation in highly populated areas remains a challenging task for robots because of the difficulty in guaranteeing safe interactions with pedestrians in unstructured situations. In this work, we present a crowd navigation control framework that delivers continuous obstacle avoidance and post-contact control evaluated on an autonomous personal mobility vehicle. We propose evaluation metrics for accounting efficiency, controller response and crowd interactions in natural crowds. We report the results of over 110 trials in different crowd types: sparse, flows, and mixed traffic, with low- ($< 0.15\ ppsm$), mid- ($< 0.65\ ppsm$), and high- ($< 1\ ppsm$) pedestrian densities. We present comparative results between two low-level obstacle avoidance methods and a baseline of shared control. Results show a 10% drop in relative time to goal on the highest density tests, and no other efficiency metric decrease. Moreover, autonomous navigation showed to be comparable to shared-control navigation with a lower relative jerk and significantly higher fluency in commands indicating high compatibility with the crowd. We conclude that the reactive controller fulfils a necessary task of fast and continuous adaptation to crowd navigation, and it should be coupled with high-level planners for environmental and situational awareness.**

***Index Terms*— Mobile Service Robots, Human-Robot Interaction, Reactive Navigation Control, People detection and tracking, autonomous navigation**

I. Introduction

Mobile service robots offer great societal value, such as transporting personal mobility devices (Segway, USA, Rokuro, Japan), last-mile delivery services (Starship Inc. USA), autonomous cleaning robots (Bluebotics, Switzerland), autonomous wheelchairs (Whill Inc. Japan), telepresence robots and tour-guide robots. Nonetheless, most robots are still limited to navigation in low-density areas, such as pathways or large open areas with the basic safety control system setting the robot to freeze as soon as it perceives a likely contact with pedestrians [1], [2]. Such a reaction would most likely be unexpected by pedestrians and lead to more dangerous collisions with pedestrians stumbling on the robot, a "frozen" robot would become a danger to itself and bystanders [3], [4]. This is the case, in highly dynamic environments such as malls, airports, markets, and mix-traffic zones with other mobility vehicles (as shown in Fig. 1).

Fig. 1. Crowd navigation evaluation around the city of Lausanne, with the robot Qolo using reactive navigation for obstacle avoidance and post-contact compliance. Top, depicts the on-board people detection and tracking from Lidar and RGBD data. Bottom, navigation view from the external camera in dense mixed crowds environment

However, guaranteeing obstacle avoidance during navigation in highly occupied areas would be unattainable for current mobile service robots bounded by actuation power, computational resources, and expected to behave as pedestrians, i.e., holonomic, reactive, communicative, and knowledgeable of proxemics and other social rules. In this work, we investigate possible reactive control strategies that avoid the "freezing" robot problem and experimentally validate the compatibility and interaction with bystander pedestrians.

Different control approaches have been proposed for dynamic obstacle avoidance, from model predictive control (MPC) applied as a quadratic program for avoiding collisions (combining braking and steering) [5], or computing interactions with other agents and formulating crowd navigation as an optimal control problem [6]. As well, the partially observable Markov decision process (POMDP) is a common framework for planning with other agents' uncertain inten-

*This work was funded in part by the EU H2020 project "Crowdbot" (779942).

† is the corresponding author. [1] D.Paez-Granados is with the SCAI Lab, Swiss Federal School of Technology in Zurich - ETH Zurich and SPZ, Switzerland dfpg@ieee.org

[2] Y. He, D. Gonon, and A. Billard are with the LASA Laboratory, Swiss Federal School of Technology in Lausanne - EPFL, Switzerland {yujie.he; david.gonon; aude.billard}@epfl.ch

[3] D. Jia, and B. Leibe are with the Visual Computing Institute, RWTH, Germany {jia; leibe}@vision.rwth-aachen.de

[4] K. Suzuki is with the Artificial Intelligence Lab, University of Tsukuba, Japan kenji@ieee.org

tions [7], [8], [9]. Other works in path planning solve the dynamic environments through models of the crowd trained in simulation. The work in [2] uses deep reinforcement learning to plan the robot's trajectory through a surrounding crowd's motion which is tracked and predicted by a separate algorithm beforehand. [10] works similarly with predictions of zones where the robot could become obstructing or freeze, and the approach then avoids these zones. While the work in [11] presents a time efficiency-based method for path planning in crowd environments. Nonetheless, none of these works guarantees continuous operation of the robot navigation or offers a solution for fully blocked passages, likely to occur in dense crowds. Moreover, none of these methods has been systematically evaluated in natural crowds.

Dynamical systems (DS) based obstacle avoidance provides a fast and continuous solution through modulation of obstacles, as offered in [12]. We have augmented this approach with a compliant control mechanism, herewith mitigating risks and allowing the robot to navigate in post-contact scenarios through sliding behaviour, as proposed in [13]. This approach offers the reactivity of time-invariant DS combined with contact estimations of impacts through a compliant bumper, herewith ensuring impact absorption through passive and active compliance for mitigating unexpected impacts with mobile robots.

In this work, we provide a framework for crowd navigation through an integrated reactive controller for obstacle avoidance and post-contact sliding control. We contribute by providing a systematic evaluation methodology and metrics for robot assessment on crowd navigation tasks. We present the results from testings in multiple natural crowd densities: low (< 0.15 $ppsm$), medium (< 0.65 $ppsm$), and high (< 1 $ppsm$), and crowd types: sparse, flows and mixed traffics.

Previous works on crowd navigation have focused on three main metrics: collisions, success rate, and time to goal [14], [11]. While the work in [11], proposed a mid-density flow evaluation ($< 0.5ppsm$) with volunteer participants in a controlled setting, and used time efficiency and boundary violations (virtual collisions) as main metrics. In contrast, we propose evaluations on natural crowds with additional metrics that assess path efficiency, controller performance, and robot-crowd interaction, similar to the simulation framework offered in [15].

We validated the whole architecture for crowd navigation and three reactive navigation controllers on a personal mobility vehicle - Qolo [16] shown in Fig. 1, a type of powered wheelchair for standing mobility of lower-limb impaired people. Although results in this work were shown on a person carrier robot, the proposed controller and metrics are equally valid for any mobile robot. We provide the whole dataset of the current experiments as open-access [1] in [17]. As well as, all source code for processing and analyzing interactions [2]. Which should enable future research on understanding people's navigation around robots, and improving detection and tracking methods.

II. Problem Statement

Obstacle avoidance methods usually consider a bi-state problem with collisions as an absolute negative state which in turn leads to the "freezing robot" problem [1], [2]. Nonetheless, contact might be inescapable when the robot's kinematic and dynamic constraints are below the pedestrians. Hence, collisions might become unavoidable even in simple scenarios. In our previous work [13], we focused on combining active compliance with DS-based obstacle avoidance which provides a way to slide around obstacles while in contact and continue moving towards the goal. Sliding in contact limits force to a determined safe threshold following safe design considerations for robot impacts [18].

In this work, we target to prove the feasibility of crowd navigation in natural crowds through reactive navigation control and developed appropriate metrics for its assessment. We compare two methods for obstacle avoidance and compared their performance with shared control (SC) as a baseline with user-given commands [19]. We explored three different controllers focusing on the following questions:

1) What difference can be observed in the representation and control of an autonomous robot around actual crowds?
2) How do the proposed reactive navigation methods perform when the crowd density changes?

A. Control Architecture for Reactive Control in Dense Crowd Navigation

In this work, we formulated the crowd navigation control by dividing it into three main layers (as depicted in Fig. 2): First, the high-level planner decides the motion direction and velocity. In the robot Qolo a user intention recognition for shared control is possible through a hands-free user interface [19], or a standard virtual joystick. Second, a reactive control layer that deals with local obstacle avoidance and post-collision control through closed-loop dynamics of the robot in its control space ($\dot{\xi}_u$). The main objective of the reactive control module is to provide a layer for local navigation assistance and immediate responsiveness to unmodeled or unpredicted object occurrences or motions. Therefore, this algorithm assumes a continuous dynamical system (DS) guiding the robot's motion to be existing and given by a high-level algorithm that optimally plans towards its intended goal (a user-given command, or a DS from planners).

For obstacle avoidance, we use a Modulated Dynamical System (MDS) [12], representing obstacles analytically as star-shaped level sets of a distance function that absorb the robot's footprint, thus, allowing the robot to be represented as a point moving in Cartesian space. In the case of Qolo, we control a single point in the bumper area as a holonomic

[1]Dataset website: https://www.epfl.ch/labs/lasa/crowdbot-dataset/

[2]Data analysis tools can be found here: https://github.com/epfl-lasa/crowdbot-evaluation-tools

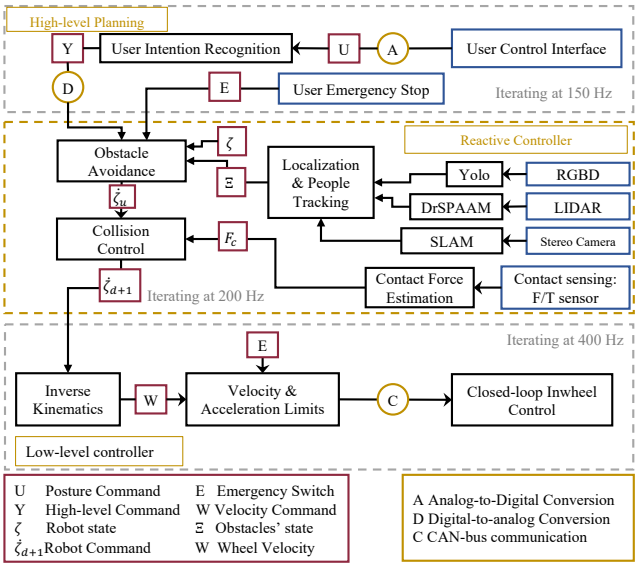


Fig. 2. Controller architecture proposed for reactive navigation control on crowds navigation. High-level planning was set from the user in shared-control low-level blending or full autonomy it was replaced by a linear DS towards an attractor at the desired location (20 to 50m ahead). Obstacle avoidance used a VO-based optimization [20] or a modulated Dynamical System (DS) [12]. Finally, a passive DS handles the post-collision through sliding [13].

point. It guarantees to lead the robot to its goal (hereafter called attractor) by assuming a circular virtual boundary (as depicted in Fig. 3). Further details of this formulation are described in [13]. Similarly, local obstacle avoidance could be achieved through the Reactive Driving Support (RDS) proposed in [20], which replicates the behaviour of the MDS locally through Velocity Obstacles (VO) based controller. While it sacrifices the guarantee to reach the global goal (by itself, i.e., without an additional path planner), it allows representing more accurately the robot shape detail and the non-holonomic constraints of the robot's kinematics.

For collision control, we have proposed a compliance and contact force controller through a sliding method using a known sensing surface over the robot's hull with a limited contact reference force F_n. Ensuring that the robot reacts to unexpected contacts (in this case, limited to the frontal bumper) and advances with a sliding manoeuvre should the underlying obstacle avoidance lead away from the contact surface without colliding with other obstacles, as we proposed in [13].

We modelled the robot dynamics as:$M\ddot{\xi} + C\dot{\xi} = \tau_c + \tau_e$, where $\dot{\xi} \in \mathbb{R}^2$ represents the robot's Cartesian velocity as a time invariant position dependent dynamical system. $M \in \mathbb{R}^{2\times2}$ corresponds to the virtual mass of the robot, $C \in \mathbb{R}^{2\times2}$ accounts for centrifugal and Coriolis terms, τ_c represents the control forces and τ_e any external disturbances to be rejected. We use a controller of the form:

$$\tau_c = \lambda_t f_u(\dot{\xi}) + (F_n + F_c)\hat{n} - D\dot{\xi} \ , \qquad (1)$$

where $f_u(\xi)$ represents the driving force generated by the nominal DS from the obstacle avoidance input, applied tangentially to the collision surface (λ_t) in case of contact. F_n represents a chosen force limit bounded by safety and acceptability, while F_c represents the estimated contact force in the $\hat{n}$ normal surface. $D \in \mathbb{R}^{2\times2}$ represents a negative defined damping effect used for controlled sliding during contacts. We transform the velocity to the domain of the robot by a first-order Taylor expansion, thus the control for the desired velocity $\dot{\xi}_d$ is,

$$\dot{\xi}_{d+1} = \frac{T_s}{M}\left((F_n + F_c)\hat{n} - D\dot{\xi}_d\right) + \hat{t}^T\dot{\xi}_u\hat{t} \ , \qquad (2)$$

where a discretizing time constant (T_s) is set to the sampling time of the control loop, and M is a virtual desired mass for impact response.

The third and final layer is a low-level control that guarantees closed-loop velocity control and state feedback to the higher-level controllers.

A key component for crowd navigation is the localization and tracking of the environment and crowd (Ξ). We have set people tracking through a real-time pipeline of sensing fusion of Lidar-based detection by DR-SPAAM [21] and RGBD detection by YOLO [22]. Whereas robot state (ξ) was estimated on optical flow principle from a stereo camera (Intel T265) fused with IMU and odometry. The full controller repository can be found in [23].

III. Evaluation Method for Crowd Navigation

A. Evaluation Metrics

Evaluation of performance in crowd navigation for mobile robots should account for the robot's performance in achieving its task, as well as the compatibility with pedestrian navigation response and safety-related metrics.

We integrate and present feasible metrics for crowd-robot interaction from embedded sensing and perception pipelines. Moreover, we make use of 3D point-cloud sensing for further assessment of the robot's navigation performance in the crowd through 3D people detection [24], and tracking [25]. The following metrics were considered:

1) Path Efficiency: **Relative time to goal**: The relative time to goal compares the time taken by the robot to reach its goal when it is alone to the time it takes when driving in the crowd as: $T_{rtg} = t_{\text{free}}\ /t_c$, where t_{free} is the free path completion time, and t_c is the robot crowd navigation time. Thus a $T_{rtg} = 1$ would be a free path or 100% efficient navigation.

Relative Path length: The relative path length compares the length of the path taken by the robot to reach its goal with and without a crowd as: $L_r = L_{goal}/L_c$, where L_{goal} is the shortest path to the goal, and L_c is the traveled path with the crowd.

Relative Jerk: The relative jerk evaluates the smoothness of the path taken by the robot to reach its goal (J_{c}) normalized to a reference jerk (J_{ref}) taken during a manual operation of the robot around the crowd as follows:

$$J_{rp} = J_{\text{c}}\ /J_{\text{ref}} \ ,$$
$$J = (1/t_f)\sum_{t=0}^{t_f}\Delta t\sqrt{J_v{}^2 + J_\theta{}^2} \ , \qquad (3)$$

where the computed jerk J takes J_v and J_θ from the linear and angular jerk, respectively. Δt corresponds to the sample time window over the period $0 \sim t_f$.

2) Controller performance: These tests aimed to estimate the level and type of assistance provided by the reactive controller towards avoiding obstacles. We adapted these metrics from shared control (SC) studies [26], [27], [28].

Contribution: C is calculated as $C = ||u_r - u_u||/||u_u||$, over a finite number of discrete samples N. Where the u_r is the user command in the robot command space (linear and angular speed), U_r is the output command of the robot. Both normalized to the maximum linear and angular velocities. In case of full autonomy with a high-level input controller, we evaluated the contribution as $C = ((||u_r||)/(||u_u||))$. Where $||x||$ represents the L2-norm of x.

Fluency: as a measurement of the commands temporal continuity given that the reactive navigation is intervening with the main driving DS desired motion as:

$$F = \frac{1}{N}\sum_{t=t_0}^{t_N} 1 - \left|u_h^t - u_h^{t-1}\right|, \tag{4}$$

where u_h is the high-level input from the user.

Agreement: We defined in terms of the deviation of the direction of the high-level commands from the direction of the final control's velocity, as follows:

$$\begin{aligned} \theta(u) &= \tan^{-1}\left(\frac{v}{w}\right), \\ a_i &= 1 - \left|\theta\left(z_u^i\right) \ominus \theta\left(u_{SC}^i\right)\right|/\pi, \\ \text{agreement} &= \sum_{i=0}^{N} a_i \cdot \Delta t_i / \sum_{i=0}^{N} \Delta t_i, \end{aligned} \tag{5}$$

where v and w are the linear and angular velocities $u = [v, w]$, a_i is the normalised agreement at time step t_i and u_{SC}^i is the final output of the robot. N is the number of samples available in which data from the high-level command input z_u^i coincide in time with u_{SC}^i , and Δt_i is the duration of the high-level input command z_u^i.

3) Pedestrian interaction metrics:: **Crowd density:** We report the crowd density in terms of mean, standard deviation and max density. However, the limited field of view of the robot as it is embedded in the crowd hinders overall measurements. Therefore, we reported on the density within 2.5, 5 and 10m around the robot (as shown in Fig. 5(c)), which explains better the scenario.

Minimal distance: the minimal distance to the robot held by any pedestrian.

Virtual Collisions: number of violations to the virtual boundary set to the robot controller, as shown in Fig. 3.

Collisions: number of collisions in a scenario reported by the experimenters and post-analysis of the recordings.

B. Experimental Setup

In this work, we present an evaluation of natural crowd navigation through reactive control. We have used the robot Qolo [16] - a person carrier mobile robot for lower-limb impaired users - instrumented as shown in Fig. 3.

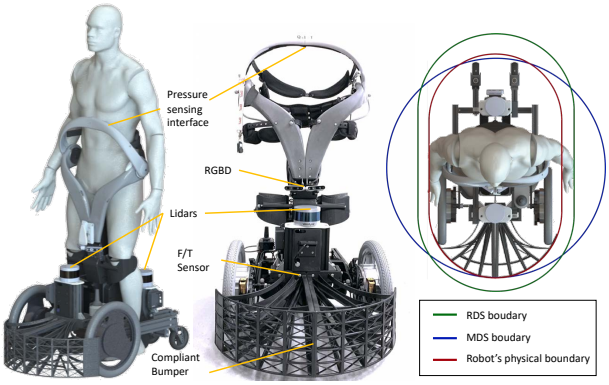


Fig. 3. Robot sensing implementation on Qolo [16]: obstacle perception with 2 LIDARs (Velodyne VLP-16), RGBD sensors (Intel Realsense D435), force sensing with one Rokubi 2.0 (Botasys). With the perception and control implemented on two embedded computers Upboard Squared, and one Nvidia Jetson Xavier AGX. On the right, the boundary representations used around the actual robot

We present a comparative evaluation of two modes of obstacle avoidance (MDS and RDS) driven by a linear DS towards a set attractor (or goal some distance from the starting point). And compare it to a baseline of shared-control where a user guides the high-level interaction with the crowd (effectively, deciding the direction of motion). Such evaluation allows observing the compatibility of the methods with the crowd, the high-level commands, and the effectiveness of pure reactive navigation. Two navigation scenarios were developed for different crowd densities:

1) Scenario 1: In this scenario, we performed crowd navigation on a mixed-traffic street (Rue de Saint-Laurent, Lausanne). The path was chosen at an intersection of 6 streets (as shown in Fig. 4, top). We encountered flows from low to mid-density with mixed crowd types: 2D flows and sparse crowds with static pedestrians. The sparse crowds' mean density was between 0.05 $ppsm$ to 0.15 $ppsm$ (40 records 50 m round trips, with 33 successful trajectories) and a max crowd density of 0.7 $ppsm$. The experiment was repeated over multiple dates during a farmer's market to elicit similar crowds.

2) Scenario 2: In this scenario, we performed 1D flow navigation during the Christmas market (Place de l'Europe, Lausanne), as depicted in Fig. 4, bottom. Here, we found denser mixed crowds formed by people lining up and flows of pedestrians. In this case, we followed the same protocol of traversing a set path whose starting and ending goals were 30 m apart.

All data were recorded on ROS standard data type (rosbag) with the following information:

- Pedestrian motion information in the form of a set of two sets of 3D point clouds around the robot, including all surrounding people and obstacles in a range of up to 50 m.
- Pedestrian's motion data from a forward-looking RGBD camera, with people, labelled and blurred.
- Output from 3 people detection layers and 1 integrated people tracker.
- Force/Torque information gathered by the contact sensors at the robot's bumper.

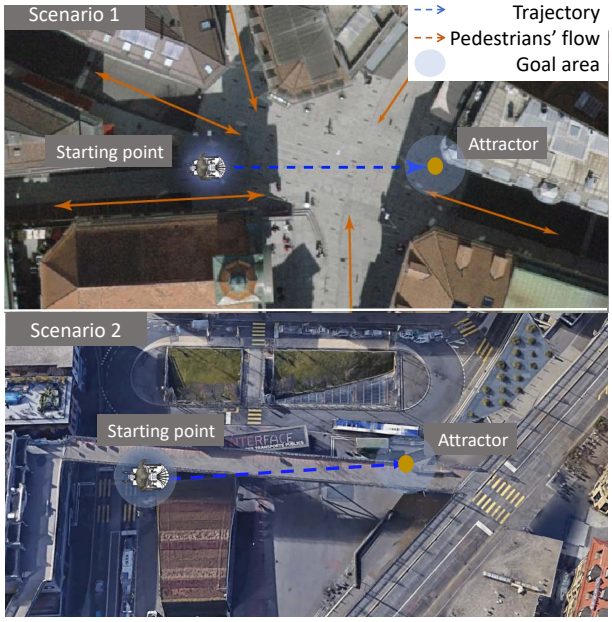


Fig. 4. Experimental setup on two mixed traffic streets with a fixed starting point and goal (attractor). Scenario 1, constitutes a mixed influx of 6 streets during a farmer's market with up to 0.7 $ppsm$. Scenario 2, is a closed pedestrian street with crowd crowds up to 1 $ppsm$, and the corresponding experiment was conducted during the Christmas market.

- Recordings of the navigation interface input given by the user/driver of the robot.
- Motion data was gathered from the robot inertia sensors and odometry sensors.
- Video recording of the scene from the robot's perspective without personal identification recording.
- *Video recording of the participant-robot driving in a scenario, with no personal identification.

All data captured have been released as an open-access dataset [17], with rosbag files and post-processing data for each of the following scenarios.

IV. Results

Scenario 1 was successfully recorded 95 times for about 5.0 km in sets of 20 m round trips (with 3 interrupted trials). A sample data is given in Fig.5 around low and mid crowd densities. Scenario 2 was successfully recorded 15 times with the highest densities close to 1 $ppsm$ from on-board measurements.

A. Controller Response in Crowds

The overall crowd density (as shown in Fig. 5(c)) measured from the onboard sensing fluctuates around 0.12 $ppsm$, with peaks of up to 0.6 $ppsm$. Although it is worth noting that the current measurements are limited to the proximity sensors, therefore, it could be possible that higher densities were not visible from the robot's viewpoint. Out of successfully recorded 16 MDS trials, 28 RDS trials, and 48 SC trials. We qualified all data of successfully reaching the goal with a margin of 3 m around the goal (as depicted in Fig. 4). Moreover, we took compatible trials with no significant difference in crowd density variance among tests (mean and max). Resulting on a set of 16 MDS trials, 20 RDS trials, and 16 SC trials for comparison.

In the interaction with the crowd, the minimal distance to pedestrians (Table I) during the trials showed to be lower in the case of the RDS controller, compared with MDS ($p < 0.05$, $F = 3.9$), which is expected for the RDS controller having a more tight definition of the robot's shape. Nonetheless, in shared control where the obstacle avoidance is provided by RDS, there was a significantly higher minimal distance ($p < 0.05$, $F = 4.4$) to surrounding pedestrians. This suggests that the high-level input from a user would prefer a higher proxemic distance to the pedestrians than the reactive controllers provided ($\sim 0.5m$).

The path efficiency results show a relative time to the goal similar among controllers, with relative times from 30% up to 60% the efficiency of a free path. No significant difference was observed among the methods. This suggests that the reactive navigation for obstacle avoidance does not hinder performance when compared with SC where the user can take lead over the direction of motion. The relative path length was similar among all recordings of MDS ($21.5 \pm 5.9m$), RDS (20.2 ± 6.2), and SC (27.1 ± 12.8), showing no significant difference.

The relative jerk (Fig. 6(d)) showed significant difference between MDS and RDS controllers ($p < 0.1, F = 7.3$), with RDS being 5% lower. RDS and SC mode ($p < 0.01, F = 41.8$), with RDS being 6% lower. On the contrary, comparing SC and MDS did not show a significant difference.

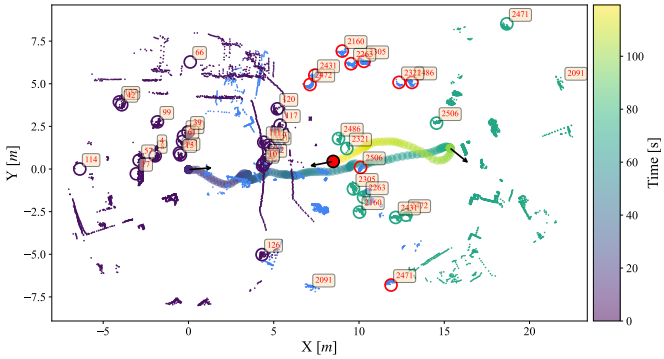
In turn, the fluency of the controllers (Fig. 6(c)) was significantly different among all tested controllers ($p < 0.01$, $F = 28.4$), however, only in SC tests was observed a noticeable difference with a 2% drop in the fluency of the user-robot commands. Overall, this high fluency demonstrates the compatibility of the reactive controllers with the high-level planner.

Analyzing the controller contribution and agreement we observed that RDS control guides most of the linear velocity components with an average 87% contribution, whereas the MDS controller provided 60% of contribution (significantly different at the level $p < 0.01$, $F = 80.6$) and agreement on the 70%. The difference in the approach to avoiding obstacles is clear, with RDS controlling the magnitude of the velocity, whereas MDS executes higher control over the angular component.

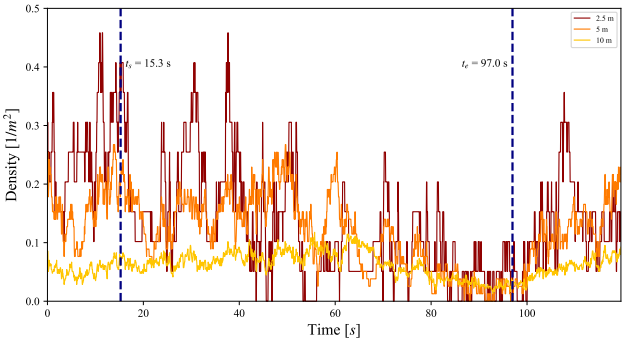
Compared with SC, the results (see, Fig. 6) show that the obstacle avoidance contributed to 50%, with agreement over 85%, which means that the user was in control of the motion direction 50% of the time, and the obstacle avoidance assisted mostly in the velocity magnitude over the tests.

1) Collisions:: As detailed in table I, we observed, across all tests, seven collisions leading to contact with pedestrians. All of which occurred at the frontal bumper of the robot. In 5 of the collisions, either the operator on-board the robot or the external supervisor of the tests pressed the emergency stop, therefore no post-collision control was recorded on those occasions. In the remaining 2 contacts (contacts with the RDS model) the contact was with a pedestrian's shopping trolley, and the contact was below impact threshold $F_c = 45N$, so that, there was not any perceived reaction from the

(a) Snapshot of the experiment where a user onboard the robot Qolo acts as safety operator while the robot drives in full autonomous mode to navigate 20 m ahead in the mixed crowd flow with parts of the trajectory in low-density $< 0.2ppsm$ ($t = 2, 12, 38$), and mid-density $< 0.7ppsm$ ($t = 24$).

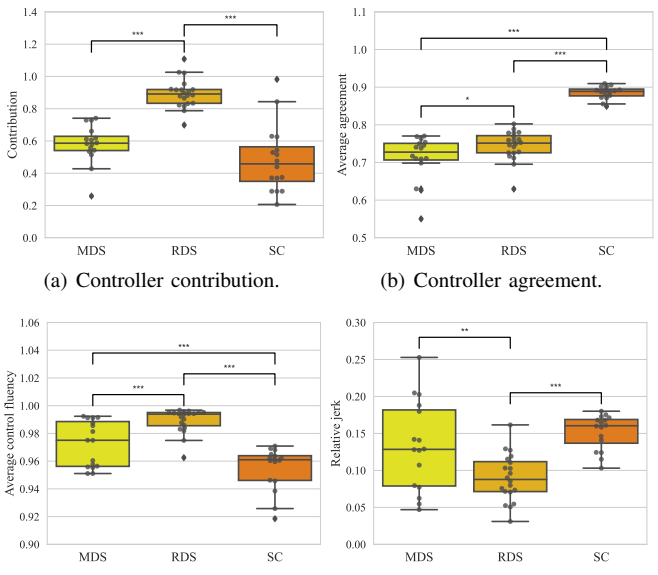


(b) Robot trajectory over one round trip to the goal location (further point to the right). Detected pedestrians are over-imposed with labelled numbers and corresponding colours at t=5, t=75, and t=118.6.



(c) Crowd density estimated from online sensing of people detection [21] at 2.5 m, 5.0 m, and 10 m around the robot. The time labels denote the start and end of the first trajectory to reach the goal. The subsequent time is the return to the starting point.

Fig. 5. Experimental setup example of one of the 110 trials of point-to-point navigation running around crowds in the city of Lausanne (Switzerland)).



(a) Controller contribution. (b) Controller agreement.

(c) Controller fluency. (d) Relative Jerk.

Fig. 6. Controllers evaluation among the task on scenario 1 on crowds between low-mid densities (up to 0.6 ppsm). All tests were compared by a one-way ANOVA at the levels $* \rightarrow p < 0.1$, $** \rightarrow p < 0.05$, and $*** \rightarrow p < 0.01$.

compliance controller, nor was collision perceived by the pedestrian.

We documented each collision case and found the following possible causes for the controller behaviour. First, limited acceleration on the robot that could not match the pedestrian's acceleration. Probable cause of one collision with RDS test, and one with SC. Second, non-holonomic

TABLE I

CONTROLLER COMPARISON IN MID-DENSITY CROWDS

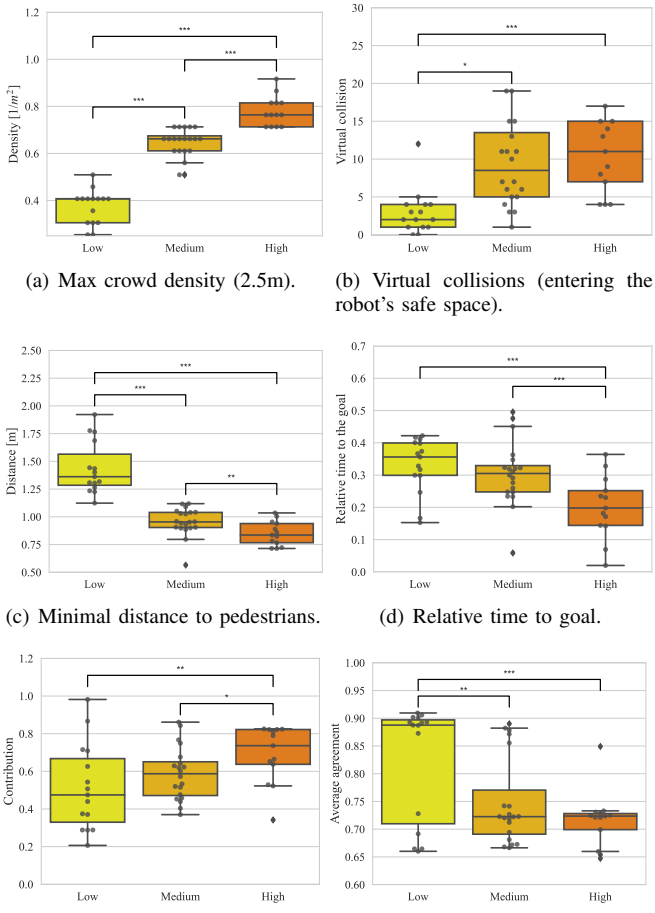
Metrics	Controller		
	MDS	**RDS**	**Shared control**
Avg. crowd density	0.12 ± 0.03	0.13 ± 0.03	0.12 ± 0.03
Max crowd density	0.45 ± 0.08	0.47 ± 0.12	0.51 ± 0.14
Min distance	1.19 ± 0.16	1.08 ± 0.18	1.20 ± 0.16
Time to goal	0.28 ± 0.09	0.32 ± 0.10	0.29 ± 0.07
Path length	1.41 ± 0.21	1.34 ± 0.20	1.52 ± 0.52
Jerk	0.13 ± 0.06	0.09 ± 0.03	0.15 ± 0.02
Contribution	0.58 ± 0.12	0.89 ± 0.09	0.49 ± 0.21
Avg. fluency	0.97 ± 0.02	0.99 ± 0.01	0.95 ± 0.02
Avg. agreement	0.71 ± 0.06	0.74 ± 0.04	0.89 ± 0.02
Virtual collision	3.50 ± 2.71	7.05 ± 7.92	4.25 ± 3.11
Actual collision	2/16	2/20	3/16

constraints in dense crowds limit reactivity, a probable cause for one collision with MDS. Third, pedestrians being very close to a surrounding obstacle generated false negatives in pedestrian detection. Probable cause for two collisions with SC and MDS tests.

Virtual collisions, defined as entering the virtual space of the robot controller (as depicted in Fig. 3). We found a slightly lower number of cases on the MDS controller compared with RDS ($p < 0.1$, $F = 2.9$), with no significant difference to SC. This result is expected given the larger size of the robot representation used in MDS, which forces the robot to be further away from pedestrians.

B. Crowd Density Influence on Controller Response

We took 48 samples from scenarios 1 and 2 of successful crowd navigation within the set points using the shared control navigation. Subsequently, we cluster the data in 3 groups



(a) Max crowd density (2.5m).

(b) Virtual collisions (entering the robot's safe space).

(c) Minimal distance to pedestrians.

(d) Relative time to goal.

(e) Controller contribution.

(f) Controller agreement.

Fig. 7. Navigation evaluation among the task on scenarios 1 and 2 comparing reactive navigation performance by crowd density. The cluster size was 15, 20, and 13 for low, medium and high density, respectively. All tests were compared by a one-way ANOVA at the levels $* \rightarrow p < 0.1$, $** \rightarrow p < 0.05$, and $*** \rightarrow p < 0.01$.

of low, medium and high densities considering average, variance and max densities in $2.5m$ and $5m$, resulting in the metrics computed in table II. The metrics with a significant difference among the density conditions are shown in Fig. 7.

The proximity of pedestrians clearly drops consistently with the crowd density ($p < 0.01$, $F = 53.3$) as shown in Fig. 7(c). Meanwhile, the average number of virtual collisions increases significantly only to medium densities ($p < 0.01$, $F = 15.3$). We observe, in sparse crowds a mean of 2.9 virtual collisions per test, whereas an average of 12 collisions per test in dense crowds.

Nonetheless, it is worth noting that in high-density crowds no actual collisions or contacts were recorded. Contrary to a medium density where we recorded 3 contacts.

Comparing the path efficiency, we found a significant difference ($p < 0.01$, $F = 8.5$) in the relative time to goal with navigation on high density dropping 10% of the relative time. In contrast, relative path length showed no significant difference among the tested crowd environments. Hence, we observe an expected behaviour compatible with human crowds where the reactive controller operates by adjusting its velocity to the crowd.

Fig. 7(e) shows the comparison of the controller contribution for the navigation. We observe no significant difference between low-mid densities, with only high density showing a significantly higher contribution ($p < 0.05$, $F = 3.9$) with a mean of 69%. In the agreement metric with the user, we observe significantly high values (82%) in low density ($p < 0.05$, $F = 5.8$), which drops to 75% in medium density, and 72% in high density. i.e., sparse low-density crowd navigation required less reactive control assistance whereas in the cases of mid- and high-density required increased reactive control contribution although no decreased agreement was registered. The user and the controller likely agree on the direction due to the more restrictive available space in the flow of dense crowds

TABLE II

CROWD NAVIGATION COMPARISON IN DIFFERENT DENSITIES

Metrics	Crowd density		
	Low	**Medium**	**High**
Avg.Density (2.5m)	0.08 ± 0.02	0.18 ± 0.03	0.26 ± 0.04
Max.Density (2.5m)	0.37 ± 0.07	0.64 ± 0.06	0.78 ± 0.06
Avg.Density (5m)	0.09 ± 0.02	0.17 ± 0.02	0.19 ± 0.02
Min.Distance	1.44 ± 0.24	0.96 ± 0.13	0.85 ± 0.11
Time to goal	0.33 ± 0.09	0.30 ± 0.10	0.20 ± 0.10
Path length	1.46 ± 0.57	1.74 ± 1.28	2.89 ± 4.07
Jerk	0.14 ± 0.03	0.14 ± 0.03	0.15 ± 0.03
Contribution	0.51 ± 0.23	0.59 ± 0.14	0.69 ± 0.15
Avg. fluency	0.95 ± 0.02	0.96 ± 0.02	0.95 ± 0.01
Avg. agreement	0.82 ± 0.11	0.75 ± 0.08	0.72 ± 0.05
Virtual collision	2.93 ± 2.94	13.15 ± 19.28	12.08 ± 8.50

V. SUMMARY AND CONCLUSIONS

In the current evaluation of a reactive controller for crowd navigation, we have shown its feasible to navigate around natural pedestrian crowds demonstrating the controller's compatibility with the behaviour of bystander pedestrians. During all 110 trials in different crowd types: sparse, flows, and mixed crowds, we did not report any incidents with pedestrians, although, we recorded 7 contacts within safe limits. On the contrary, any comments made by bystanders were positive about the perceived utility of assistive navigation technology for mobility-impaired users.

The reactive navigation controllers showed to perform similarly in trajectory efficiency and fluency during the motions. The controllers differ in the way to avoid obstacles, one mainly deviating through angular velocity (MDS) and the other mostly by reducing speed (RDS). It is hence expected that they lead to measurable differences such as faster displacement to the goal for MDS than RDS (confirmed by results), larger jerkiness for MDS than RDS (confirmed by data). One may have expected that MDS leads to more collisions than RDS but this is not confirmed by the results. However, the number of collisions were too few to conclude.

A significant difference found in the reactive controller response was given by the type of crowd, where sparse and mixed traffic crowds in the medium density showed to be more difficult to navigate with a lower agreement with the user at the same level of contribution. In contrast, flow crowds made it easier for the controller to respond to a somewhat uniform response of the surrounding pedestrians, as validated with the equal agreement levels (see, Fig. 7).

The proposed reactive controller focused on a single layer of the robot response for crowd navigation which performs efficiently compared with shared control, with no significant difference in time to goal, minimal distance to pedestrians, motion jerk, and path length. Further work should focus on situational awareness of the complex mixed crowd environments, as required when interacting in complex environments, e.g., on some occasions, the robot did not interact appropriately with the environment. i.e., getting inside stores, getting in the middle of people lining up, and entering shops in the wrong direction. In comparison, shared control was more natural, and most of the tests ran smoothly without bystanders even noticing the robot.

ACKNOWLEDGEMENT

The experiments were approved with an ethical protocol by the human research ethical committee of EPFL (Approval No: HREC-032-2019), and with the approval and cooperation of the office for mobility, the police and the office for parks, and the public domain of the city of Lausanne. Approval numbers: 395128 and 416008. *Disclaimer:* D.P. and K.S. hold the patents of the robot Qolo and shares in the company Qolo Inc.

REFERENCES

[1] P. Trautman and K. Patel, "Real time crowd navigation from first principles of probability theory," *Proceedings of the International Conference on Automated Planning and Scheduling*, vol. 30, no. 1, pp. 459–467, Jun. 2020.

[2] A. J. Sathyamoorthy, J. Liang, U. Patel, T. Guan, R. Chandra, and D. Manocha, "Densecavoid: Real-time navigation in dense crowds using anticipatory behaviors," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 11 345–11 352.

[3] P. Salvini, D. Paez-Granados, and A. Billard, "On the Safety of Mobile Robots Serving in Public Spaces: Identifying gaps in en ISO 13482:2014 and calling for a new standard," *ACM Transactions on Human-Robot Interaction*, vol. 10, no. 3, pp. 1–29, 2021. [Online]. Available: https://doi.org/10.1145/3442678

[4] ——, "Safety Concerns Emerging from Robots Navigating in Crowded Pedestrian Areas," *International Journal of Social Robotics (SORO)*, vol. 14, no. 2, pp. 441–462, 2022. [Online]. Available: https://doi.org/10.1007/s12369-021-00796-4

[5] B. Yi, S. Gottschling, J. Ferdinand, N. Simm, F. Bonarens, and C. Stiller, "Real time integrated vehicle dynamics control and trajectory planning with mpc for critical maneuvers," in *2016 IEEE Intelligent Vehicles Symposium (IV)*, 2016, pp. 584–589.

[6] Y. Chen, F. Zhao, and Y. Lou, "Interactive model predictive control for robot navigation in dense crowds," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, pp. 1–13, 2021.

[7] H. Bai, S. Cai, N. Ye, D. Hsu, and W. S. Lee, "Intention-aware online pomdp planning for autonomous driving in a crowd," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, 2015, pp. 454–460.

[8] Y. Luo, P. Cai, A. Bera, D. Hsu, W. S. Lee, and D. Manocha, "Porca: Modeling and planning for autonomous driving among many pedestrians," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 3418–3425, 2018.

[9] P. Cai, Y. Luo, D. Hsu, and W. S. Lee, "Hyp-despot: A hybrid parallel algorithm for online planning under uncertainty," *The International Journal of Robotics Research*, vol. 40, no. 2-3, pp. 558–573, 2021.

[10] A. J. Sathyamoorthy, U. Patel, T. Guan, and D. Manocha, "Frozone: Freezing-free, pedestrian-friendly navigation in human crowds," *IEEE Robotics and Automation Letters*, vol. 5, no. 3, pp. 4352–4359, 2020.

[11] Y. Kobayashi, T. Sugimoto, K. Tanaka, Y. Shimomura, F. J. Arjonilla Garcia, C. H. Kim, H. Yabushita, and T. Toda, "Robot Navigation Based on Predicting of Human Interaction and its Reproducible Evaluation in a Densely Crowded Environment," *International Journal of Social Robotics*, 2021.

[12] L. Huber, A. Billard, and J.-J. Slotine, "Avoidance of Convex and Concave Obstacles With Convergence Ensured Through Contraction," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 1462–1469, 2019. [Online]. Available: https://doi.org/10.1109/lra.2019.2893676

[13] D. Paez-Granados, V. Gupta, and A. Billard, "Unfreezing social navigation : Dynamical systems based compliance for contact control in robot navigation," in *IEEE International Conference on Robotics and Automation (ICRA-2022)*, vol. 1, 5 2022, pp. 8368–8374. [Online]. Available: https://doi.org/10.1109/ICRA46639.2022.9811772

[14] C. Chen, Y. Liu, S. Kreiss, and A. Alahi, "Crowd-robot interaction: Crowd-aware robot navigation with attention-based deep reinforcement learning," in *IEEE International Conference on Robotics and Automation*, 2019.

[15] F. Grzeskowiak, D. Gonon, D. Dugas, D. Paez-Granados, J. Chung, J. Nieto, R. Siegwart, A. Billard, M. Babel, and J. Pettré, "Crowd against the machine: A simulation-based benchmark tool to evaluate and compare robot capabilities to navigate a human crowd," in *IEEE International Conference on Robotics and Automation (ICRA-2021)*, 2021. [Online]. Available: https://doi.org/10.1109/ICRA48506.2021.9561694

[16] D. Paez-Granados, H. Kadone, M. Hassan, Y. Chen, and K. Suzuki, "Personal mobility with synchronous trunk-knee passive exoskeleton: Optimizing human-robot energy transfer," *IEEE/ASME Transactions on Mechatronics*, vol. 1, pp. 1–12, 2022. [Online]. Available: https://doi.org/10.1109/TMECH.2021.3135453

[17] D. Paez-Granados, Y. He, D. Gonon, L. Huber, and A. Billard, "3d point cloud and rgbd of pedestrians in robot crowd navigation: Detection and tracking," *IEEE DataPort*, 12 2021. [Online]. Available: https://doi.org/10.21227/ak77-d722

[18] D. Paez-Granados and A. Billard, "Crash test-based assessment of injury risks for adults and children when colliding with personal mobility devices and service robots," *Nature Scientific Reports*, vol. 12, pp. 1–13, 3 2022. [Online]. Available: https://doi.org/10.1038/s41598-022-09349-9

[19] Y. Chen, D. Paez-Granados, H. Kadone, and K. Suzuki, "Control interface for hands-free navigation of standing mobility vehicles based on upper-body natural movements," in *IEEE International Conference on Intelligent Robots and Systems*. IEEE, 2020, pp. 11 322–11 329. [Online]. Available: https://doi.org/10.1109/IROS45743.2020.9340875

[20] D. J. Gonon, D. Paez-Granados, and A. Billard, "Reactive Navigation in Crowds for Non-holonomic Robots with Convex Bounding Shape," *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 4728–4735, 2021. [Online]. Available: https://doi.org/10.1109/LRA.2021.3068660

[21] D. Jia, A. Hermans, and B. Leibe, "DR-SPAAM: A Spatial-Attention and Auto-regressive Model for Person Detection in 2D Range Data," in *International Conference on Intelligent Robots and Systems (IROS)*, 2020.

[22] G. Jocher, Y. Kwon, guigarfr, perry0418, J. Veitch-Michaelis, Ttayu, D. Suess, F. Baltacı, G. Bianconi, IlyaOvodov, Marc, e96031413, C. Lee, D. Kendall, Falak, F. Reveriano, FuLin, GoogleWiki, J. Nataprawira, J. Hu, LinCoce, LukeAI, NanoCode012, NirZarrabi, O. Reda, P. Skalski, SergioSanchezMontesUAM, S. Song, T. Havlik, and T. M. Shead, "ultralytics/yolov3: v9.5.0 - YOLOv5 v5.0 release compatibility update for YOLOv3," Apr. 2021. [Online]. Available: https://doi.org/10.5281/zenodo.4681234

[23] D. Paez-Granados and V. Gupta, "ROS-based Qolo's controller architecture for shared control and crowd navigation," 3 2022. [Online]. Available: https://github.com/DrDiegoPaez/qolo_ros/

[24] D. Jia and B. Leibe, "Person-MinkUNet: 3D Person Detection with LiDAR Point Cloud," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2021.

[25] X. Weng, J. Wang, D. Held, and K. Kitani, "3D Multi-Object Tracking: A Baseline and New Evaluation Metrics," *IROS*, 2020.

[26] T. Carlson and Y. Demiris, "Collaborative control for a robotic wheelchair: Evaluation of performance, attention, and workload," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 42, pp. 876–888, 2012.

[27] D. Gopinath, S. Jain, and B. D. Argall, "Human-in-the-loop optimization of shared autonomy in assistive robotics," *IEEE Robotics and Automation Letters*, vol. 2, pp. 247–254, 2016.

[28] A. Erdogan and B. D. Argall, "The effect of robotic wheelchair control paradigm and interface on user performance, effort and preference: An experimental assessment," *Robotics and Autonomous Systems*, vol. 94, pp. 282–297, 2017.