

Cooperative Navigation for Mixed Human-Robot Teams Using Haptic Feedback

Stefano Scheggi, *Member, IEEE*, Marco Aggravi, *Student Member, IEEE*,
and Domenico Prattichizzo, *Fellow, IEEE*

Abstract—In this paper, we present a novel cooperative navigation control for human-robot teams. Assuming that a human wants to reach a final location in a large environment with the help of a mobile robot, the robot must steer the human from the initial to the target position. The challenges posed by cooperative human-robot navigation are typically addressed by using haptic feedback via physical interaction. In contrast to that, in this paper we describe a different approach, in which the human-robot interaction is achieved via wearable vibrotactile armbands. In the proposed work the subject is free to decide her/his own pace. A warning vibrational signal is generated by the haptic armbands when a large deviation with respect to the desired pose is detected by the robot. The proposed method has been evaluated in a large indoor environment, where fifteen blindfolded human subjects were asked to follow the haptic cues provided by the robot. The participants had to reach a target area, while avoiding static and dynamic obstacles. Experimental results revealed that the blindfolded subjects were able to avoid the obstacles and safely reach the target in all of the performed trials. A comparison is provided between the results obtained with blindfolded users and experiments performed with sighted people.

Index Terms—Human-robot team, Human-robot interaction, Haptic feedback, Formation control, Autonomous vehicles, Human body tracking, Psychophysics.

I. INTRODUCTION

LET us assume that a human wants to reach a location in a large environment with the help of a mobile robot (Fig. 1). Such situation may involve: (i) assisting an elderly or a visually-impaired person; (ii) helping a person who is in a dangerous situation with poor visibility, and hearing severely reduced due to environmental noise; (iii) human-robot cooperative tasks, e.g., holding and transporting a heavy and/or large object where the human’s pose should be corrected to avoid robot singularities and improve task performance. In our approach, the human is free to select the desired walking speed, and the robot does not force her/him to its pace as long as environmental obstacles are avoided and she/he is able to safely reach the target location. The robot guides the human only by adjusting her/his heading, in a way that the person

The research leading to these results has received funding from the European Union Seventh Framework Programme FP7/2007-2013 under grant agreement n. 601165 of the project “WEARHAP - WEARable HAPtics for humans and robots” and from the European Union’s Horizon 2020 research and innovation programme - Societal Challenge 1 (DG CONNECT/H) under grant agreement n. 643644 of the project “ACANTO: A CyberphysicAI social NeTwOrk using robot friends”.

Stefano Scheggi, Marco Aggravi, and Domenico Prattichizzo are with the Department of Information Engineering and Mathematics, University of Siena, Via Roma 56, I-53100 Siena, Italy. Domenico Prattichizzo is also with the Department of Advanced Robotics, Istituto Italiano di Tecnologia, via Morego 30, I-16163 Genova, Italy. {scheggi,aggravi,prattichizzo}@dii.unisi.it

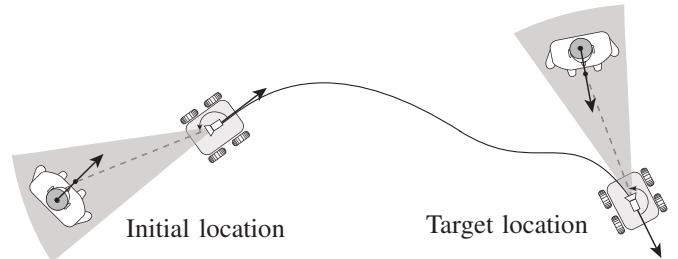


Fig. 1. Cooperative human-robot navigation from an initial to a target location (top view). The goal is to guide a human in a large environment with the help of a mobile robot. The human is free to select the walking speed. The interaction between the user and the robot is obtained via wearable haptic interfaces. They provide the user with directional cues in order to reach the target, while avoiding static and dynamic obstacles. The robot and the human are respectively equipped with a *vision sensor* and *vibrotactile armbands*. The field of view of the vision sensor is shaded.

always remains in charge of the final decision to take. The user can always override the *suggestions* given by the system. The type of correction provided by the robot has to be perceived as very *soft*, and unnatural stimulations must be avoided as much as possible.

In this work, we use haptic signals provided by tactile devices to correct the human’s pose. In real world scenarios, visual and auditory perception may be overloaded with information, thus resulting in a rapid error increase and in an overall reduction of user performance, if directional cues are provided through these channels. A possible solution is to deliver this information through an underutilized sense, i.e., the sense of touch. Similarly to sound, a tactile stimulus is made up of a signal with varying frequency and amplitude. Different from the auditory feedback, tactile stimuli directly engage human motor learning system [1] with extraordinary sensitivity and speed [2]. Moreover, tactile communication can be used in situations where visual or auditory stimuli are distracting, impractical, unavailable or unsafe.

The main source of inspiration for this work came from [3], where a passive approach inspired by the classical “Cobot” philosophy [4] was adopted for guiding an elderly person using the brakes of a commercial walker, and from [5] where the authors proposed a leader-follower formation control strategy, which in this paper has been adapted to our particular human-robot setup.

A large body of literature exists on the theme of *assistive robotics* and *human-robot cooperation/navigation*. Strictly related to our work is the study presented in [6]. The authors investigated the design of a stiff handle to enhance human trust

and confidence in cooperative human-robot tasks. Their final design consisted of a rigid handle attached to a mobile robot via a mechanical feedback spring system at the base. When the user was aligned with the robot, the spring system had zero tension. When the handle was rotated, the spring system introduced tension on the device, which increased with the rotation angle. In [7], the authors developed an assistive-guide robot to help visually-impaired people to navigate through unfamiliar areas. The human-robot interaction was achieved using a leash. In these works, the robot control did not take into account the motion of the user. Moreover, the authors did not focus on the way the human had interacted with the robot. In [8], [9], the authors presented the identification of the human-human interactions via a rigid handle along a given path. The derived interaction model can be used to design human-robot guidance policies for helping people move in dangerous situations, where they cannot use their principal sensory modalities.

In the aforementioned works, the human-robot interaction was achieved via kinesthetic feedback (haptic feedback perceived by sensors in the muscles, joints, tendons). While kinesthetic feedback is common in haptic systems, in this work we use vibrotactile interfaces. The motivation is that tactile devices are generally more portable, less encumbering and have a wider range of action than the kinesthetic ones [10]. Different from the works mentioned above, our approach does not require physical interaction between the human and the robot. In fact, although the kinesthetic feedback can be used to guide the human subject in a more effective way than by using vibrotactile stimuli, we consider the physical interaction limiting: (i) the user has her/his hands busy, thus other physical tasks may not be accomplishable; (ii) it is difficult to extend the physical interaction to multiple users. Moreover, since we use wearable devices, the proposed approach can be extended to other body parts. It can be combined/extended to guide the arms of the user along feasible trajectories in cooperative manipulation/transportation tasks. Different from [6], [7], we design a control policy that adjusts the linear velocity of the robot according to the walking speed of the user.

Haptic feedback for human guidance was considered in [11]. The authors used a grounded haptic manipulator to apply kinesthetic force and position signals to the user's hand, to assist the operator in reaching a desired position in large remote environments. Wearable haptic devices were used in [12], where a haptic belt was used for waypoint navigation. The system relied on vibrotactile stimuli and GPS information. A similar approach was used in [13], where the authors presented a navigation system that guided a human toward a goal area via a vibrotactile belt. Similarly to our work, they modeled the human as a nonholonomic robot. However, they used a different way to provide vibrotactile stimuli to the user. They also did not consider haptic stimuli for human-robot cooperative navigation, and they did not present a human-robot formation control algorithm. Finally, in [14], the authors proposed a mobile device for human navigation using multimodal communication (audio, visual, vibrotactile and directional skin-stretch stimuli).

For human-robot cooperation, recent studies have proved

the importance of haptic feedback for role negotiation in human-robot co-manipulation tasks (cf. [15], [16]). Similarly, in [17] the authors proposed an approach that exploited the arm compliance of a humanoid robot to follow the human guidance in a physical human-robot cooperative task.

A. Original contributions and organization

Our setup consists of a mobile robot equipped with a vision sensor, e.g., a Red-Green-Blue-Depth (RGB-D) camera, and a human subject wearing custom-design vibrotactile interfaces. In this work, we assume that the robot has a map of the environment. The robot autonomously localizes its pose within the map and guides the user along obstacle-free paths. Obstacle-free paths are computed for both the robot and the user and updated as soon as new obstacles are detected by the robot. Since a predefined distance and orientation should be maintained between the human and the robot at all times, the leader-follower formation control strategy proposed in [5] has been adapted to our human-robot setup. In fact, recent studies [18] have shown a close relationship between the shape of human locomotor paths in goal-directed movements, and the simplified kinematic model of a wheeled mobile robot. In particular, the authors have shown that the shoulders can be considered as a steering wheel that drives the human body. This observation indicates that humans mostly perform smooth paths and the direction of their body is tangent to the trajectories they perform.

In our scenario the human should always be able to freely select her/his walking speed. Nevertheless, a specific haptic feedback is sent to the user in order to adjust her/his heading according to the formation specifics. In this work, our goal is to send easily processable signals to the human (by exploiting the simplified model of her/his walking motion), so that she/he can promptly respond to the stimuli of the guiding robot.

The rest of the paper is organized as follows. Sections II and III present our control strategy and the haptic-feedback generation mechanism, respectively. Section IV describes our human visual detection algorithm and Section V reports the results of experimental validations. In Section VI conclusions are drawn and possible subjects of future research are outlined.

II. PROBLEM FORMULATION AND CONTROL DESIGN

In this section, we briefly review the leader-follower formation control strategy for nonholonomic robots. Then, we show how to adapt it to our human-robot setup. Note that, since in goal-directed paths the human can be modeled as a unicycle robot [18], leader-follower formation control can also be applied (with suitable modifications) to a mixed human-robot formation (cf. [19], [20]).

A. Leader-follower formation control for nonholonomic robots

Let us consider a robot whose kinematics can be abstracted as a unicycle model,

$$\dot{x} = v \cos \theta, \quad \dot{y} = v \sin \theta, \quad \dot{\theta} = \omega, \quad (1)$$

where $\mathbf{R} = (x, y, \theta)^T \in \mathbb{R}^2 \times \mathbb{S}^1$ is the pose of the robot and $(v, \omega)^T$ is the velocity control input. By $\mathbf{P} = (x, y)^T$ we denote the position of the robot, while θ represents its heading.

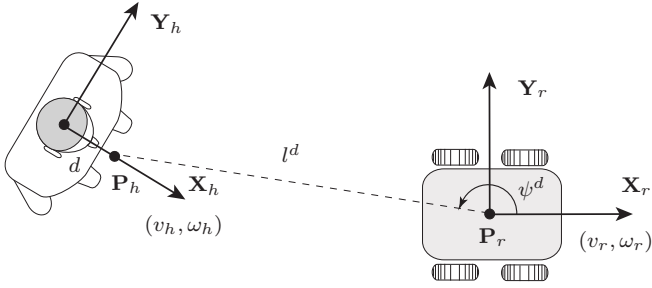


Fig. 2. Human-robot setup: l^d and ψ^d represent the desired distance and orientation between the robot and an off-axis reference point \mathbf{P}_h on the human with offset d . The human and the robot move with linear and angular velocity $(v_h, \omega_h)^T$ and $(v_r, \omega_r)^T$, respectively.

With these definitions at hand, let us briefly review the leader-follower formation control for unicycles [5]. In [5], robot \mathbf{R}_h (in our framework a *human*) must follow the robot \mathbf{R}_r with a desired separation l^d and desired relative bearing ψ^d (Fig. 2). The formation control problem consists of determining the velocities $(v_h, \omega_h)^T$ of the follower, which maintain the formation as the leader moves. Let $\beta = \theta_r - \theta_h$ be the relative orientation of \mathbf{R}_r and \mathbf{R}_h , $\mathbf{u}_r = (v_r, \omega_r)^T$ and $\mathbf{u}_h = (v_h, \omega_h)^T$ their velocity control inputs, and

$$\mathbf{G} = \begin{bmatrix} \cos \gamma & d \sin \gamma \\ -\frac{\sin \gamma}{l} & \frac{d \cos \gamma}{l} \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} -\cos \psi & 0 \\ \frac{\sin \psi}{l} & -1 \end{bmatrix},$$

where d is the offset to an off-axis reference point \mathbf{P}_h on \mathbf{R}_h , $\gamma = \beta + \psi$ and l, ψ are the actual separation and relative bearing of \mathbf{R}_h and \mathbf{R}_r , respectively (Fig. 2). The desired formation velocity for \mathbf{R}_h can then be written as,

$$\mathbf{u}_h = \mathbf{G}^{-1}(\mathbf{q} - \mathbf{F} \mathbf{u}_r), \quad (2)$$

being \mathbf{q} an auxiliary control input defined as

$$\mathbf{q} = \begin{bmatrix} k_1(l^d - l) \\ k_2(\psi^d - \psi) \end{bmatrix},$$

where k_1, k_2 are positive control gains (observe that \mathbf{G} is always invertible as long as $d/l > 0$, which is always true). Equation (2) has been obtained by applying input-output linearization [21].

In what follows, we show how to tailor (2) to our human-guidance problem. Notice that in our framework the distinction between the leader and the follower vanishes: in fact, both agents cooperate to achieve a common goal (reach the desired target), without direct physical interaction.

B. Human-robot guidance

Different from (2), in our scenario the human should always be able to freely choose her/his walking speed (linear velocity v_h). However, in order to be guided by the robot \mathbf{R}_r towards a target position, her/his angular velocity ω_h should be regulated. Note that by changing the angular velocity ω_h of the user, we modify her/his heading θ_h (cf. (1)). Nevertheless, the robot should regulate its linear velocity v_r accordingly to the user, while its angular velocity ω_r depends on the specific trajectory from the initial to the target position. We assume

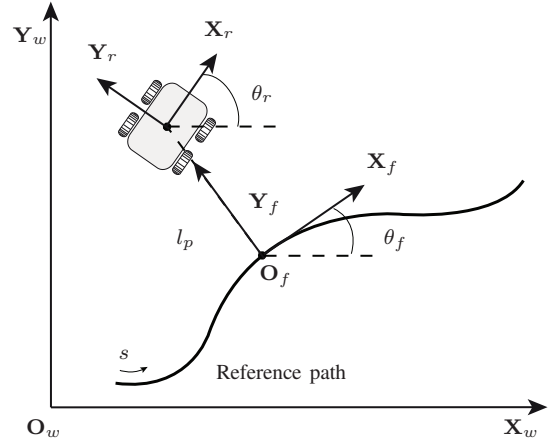


Fig. 3. Path following setup: l_p represents the coordinate of the vehicle position along the y -axis of the Frenet frame $\langle \mathbf{O}_f, \mathbf{X}_f, \mathbf{Y}_f \rangle$, s is the curvilinear coordinate of the robot along the path, θ_f and θ_r represent the angle between the x -axis of the world frame $\langle \mathbf{O}_w, \mathbf{X}_w, \mathbf{Y}_w \rangle$ and the x -axis of the Frenet and robot frame, respectively.

that the trajectory of the robot is smooth (its tangent is well-defined at each point), and its curvature is known at all points. Moreover, we consider that the vehicle is always localized with respect to the path, and that a Frenet frame, whose origin is the orthogonal projection of the vehicle position on the path, is always available (Fig. 3). If we assume that the initial robot configuration is not far from the desired path and that $v_r > 0$, the desired angular velocity ω_r of the robot that solves the path following problem is

$$\omega_r = v_r a_r, \quad (3)$$

being

$$a_r = -k_3 l_p \frac{\sin(\theta_p)}{\theta_p} - k_4 \theta_p + \cos(\theta_p) \frac{c(s)}{1 - c(s)l_p},$$

where k_3, k_4 are positive control gains, l_p represents the signed distance of the vehicle position along the y -axis of the Frenet frame, $\theta_p = \theta_r - \theta_f$, being θ_f the angle between the x -axis of the world frame $\langle \mathbf{O}_w, \mathbf{X}_w, \mathbf{Y}_w \rangle$ and the x -axis of the Frenet frame, s is the curvilinear coordinate along the path, and $c(s)$ is the curvature of the path at that point, defined as $c(s) = d\theta_f/ds$ [22].

Concerning the cooperative navigation control law, let

$$\mathbf{G}_f = \begin{bmatrix} d \sin \gamma & -\cos \psi \\ \frac{d \cos \gamma}{l} & \frac{\sin \psi}{l} \end{bmatrix}, \quad \mathbf{F}_f = \begin{bmatrix} \cos \gamma & 0 \\ -\frac{\sin \gamma}{l} & -1 \end{bmatrix},$$

then the control velocities for the human and robot are given by (cf. (2)),

$$\begin{bmatrix} \omega_h \\ v_r \end{bmatrix} = (\mathbf{G}_f)^{-1} \left(\mathbf{q} - \mathbf{F}_f \begin{bmatrix} v_h \\ \omega_r \end{bmatrix} \right). \quad (4)$$

It is worth noting that input-output linearization is possible as long as $\cos(\gamma - \psi) \neq 0$. Assuming that the human is moving with linear velocity v_h and the robot is rotating with angular velocity ω_r , then the control law reported in (4) allows to maintain the formation specified by l^d and ψ^d .

By combining (3) and (4), we obtain the following human-robot control law, which allows the robot to follow a precomputed path,

$$\begin{bmatrix} \omega_h \\ v_r \end{bmatrix} = (\mathbf{G}_{pf})^{-1}(\mathbf{q} - v_h \mathbf{F}_{pf}), \quad (5)$$

being

$$\mathbf{G}_{pf} = \begin{bmatrix} d \sin \gamma & -\cos \psi \\ \frac{d \cos \gamma}{l} & \frac{\sin \psi}{l} - a_r \end{bmatrix}, \quad \mathbf{F}_{pf} = \begin{bmatrix} \cos \gamma \\ -\sin \gamma \end{bmatrix}.$$

Note that \mathbf{G}_{pf} is not invertible if $l^{-1} \cos(\gamma - \psi) - a_r \sin \gamma = 0$ which is equivalent to $\beta = \arccos(l a_r \sin \gamma)$.

Remark 1: Suppose that the robot estimates the human motion using an onboard vision sensor with limited field of view (FoV) (cf. Fig. 1 and Sect. IV). Since the formation parameters are fixed with respect to the robot, a proper choice of l^d and ψ^d allows to maintain the human inside the sensor's field of view. \diamond

In this section, we presented a cooperative navigation control for human-robot teams. We model the human and the robot as first order systems with velocity control inputs. Hence, suitable velocities should be provided to the agents in order to safely move in the environment. While it is simple to apply desired velocities to a robot, it is not trivial to impose a desired angular velocity to a human. In the next section, we will show how we can use haptic feedback to address this problem.

III. HAPTIC FEEDBACK

Our purpose is to provide haptic stimuli in order to adjust the heading of the user. Due to the nonholonomic nature of the human locomotion in goal directed path, the device should elicit only three basic behaviors on the human (*turn left*, *turn right*, and *slow down*). Thus, only three stimuli would be sufficient in principle. In order not to overload the tactile channel and not to reduce the recognition time, we display few significative signals. Note that, although the human is always free to decide her/his pace, the *slow down* behavior is introduced in case of emergency, danger or when the maximal linear velocity of the robot is not sufficient to keep up with the human's velocity.

Different from [12], [13], which developed a vibrotactile belt to guide the user, in this work we focus on vibrotactile armbands. We reduce the number of the vibrating motors (tactors) in order to elicit only the necessary human's behaviors. The haptic interfaces are designed in order to be informative, easy to use, and wearable. In what follows, we present the vibrotactile devices and two haptic cueing methods. The first method, which consists of a more wearable solution, is composed of a single armband worn on the dominant forearm (*unilateral* condition). The second method, which aims to be more intuitive, uses two armbands placed bilaterally on the forearms (*bilateral* condition). Even if the bilateral condition allows for a larger spatial separation between the stimuli, and a better discrimination of the directional cues, the unilateral condition provides a more compact solution.

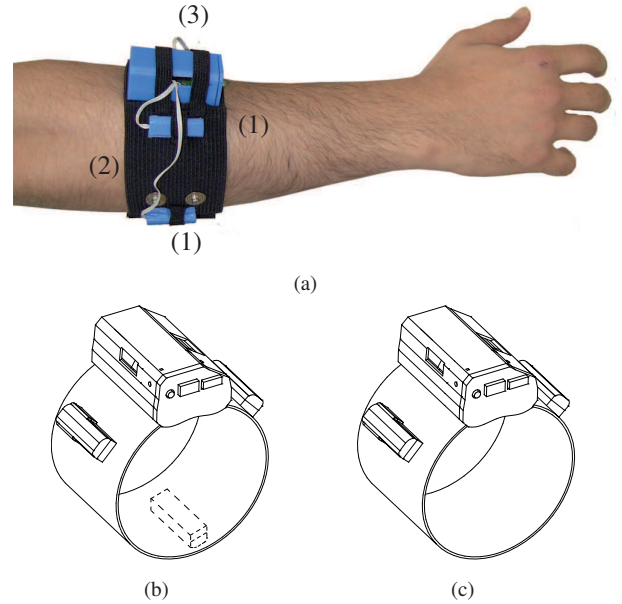


Fig. 4. (a) The vibrotactile armband is fitted on the forearm and it is equipped with vibrating motors (tactors) (1), attached to an elastic armband (2) whose width is about 60 mm. The Li-Ion battery and the Arduino board are in (3). Two different configuration were tested: (b) a single armband with three tactors; (c) two armbands with two tactors each.

A. Description of the haptic armband

Tactile vibratory sensitivity is influenced by the spatial location on the body, the distance between the stimulators, the frequency of stimulation and the age of the user. Studies have demonstrated that vibration is best detected on hairy skin due to skin thickness and nerve depth, and that vibrotactile stimuli are best detected in bony areas [23]. In particular, wrists and spine are preferred for detecting vibrations, with arms next in line [24]. Movement can decrease detection rate while increasing response time of particular body parts. For example, walking affects lower body sites the most [24]. The effect of movement on vibrotactile sensitivity has been also investigated in [25].

As with other sensory modalities, touch deteriorates with age. Discriminative capabilities and the appreciation of temporal gaps in vibratory stimuli were found to be poorer in elderly individuals. Such loss has been attributed to physiological changes in the skin itself, and/or to neurological factors. An accurate analysis of the ability to localize vibrotactile stimuli on the forearm was conducted in [26]. The authors considered the locations of the stimuli, the proximity of such stimuli to body references such as the elbow and the wrist, and the age of the users. Results showed that if points of stimulation lie adjacent to natural anchor points, stimuli localization is enhanced at those sites.

Due to the aforementioned considerations, we have concentrated on the development of vibrotactile armbands. By focusing on a single armband (unilateral condition), three tactors are sufficient to warn the user, since the haptic feedback should elicit three basic behaviors. An armband shape with three tactors circling the forearm (Fig. 4(a)-(b)) ensures sufficient distance between the vibrating motors, while covering a minimal area of the forearm. In fact, in two-point

discrimination perception, the minimal distance between two stimuli to be differentiated is about 35 mm on the forearms. There is no evidence for differences among the left and right sides of the body, and women are known to be more sensitive than men to skin stimulation [27], [23]. In order to improve the *intuitiveness* of the haptic feedback, we investigate a second solution (bilateral condition), in which two haptic armbands, equipped with two tactors each, are used. The subject wears one vibrotactile armband on each forearm in order to maximize the stimuli separation, while keeping the discrimination process as intuitive as possible. According to [26], in each modality we place the armbands close to the elbow in order to increase the separation between the tactors and exert the strongest influence on localization accuracy, due to the proximity to body landmarks.

From a technical point of view, the vibrotactile armbands are composed of cylindrical vibro-motors, independently controlled via an external PC using the Bluetooth communication protocol (Fig. 4). The communication is realized with an RN42 Bluetooth module connected to an Arduino mini pro 3.3V with a baud rate of 9600. An Atmega 328 microcontroller installed on the Arduino board is used to independently control the vibration amplitude of each motor. The Precision Microdrives 303-100 Pico Vibe 3.2 mm vibration motors are placed inside the fabric pockets on the external surface of the armband (the width of the armband is about 60 mm), with shafts aligned with the forearm (Fig. 4). The motors have a vibration frequency range of about 100-280 Hz (the maximal sensitivity of human skin is achieved around 200-300 Hz [28]), typical normalized amplitude of 0.6 G, lag time of 21 ms, rise time of 32 ms and stop time of 35 ms. Note that the proposed motors are controlled by applying a certain amount of voltage that determines both frequency and amplitude. Thus, users feel changes in both the intensity and pitch of perception when the applied voltage is varied.

B. Haptic feedback generation

In what follows, we illustrate our idea on how to convey motion information by using the proposed haptic devices. At first, we present the haptic feedback mechanism from a high level point of view. It is worth noting that the proposed mechanism is general and independent from the two vibrotactile configurations described above. Successively, we present how the haptic feedback policy is translated into vibrating stimuli for the two proposed configurations.

Let us consider three stimuli, L (*turn left*), S (*slow down*) and R (*turn right*) and let $s_j(t)$ be the vibrating signal of stimulus $j \in \{L, S, R\}$ at time t . Let Δt be the constant sampling time of our system. $\omega_h^*(t + \Delta t)$ represents the angular velocity computed by the controller in (4)-(5), and $\hat{\omega}_h(t + \Delta t)$ is the predicted angular velocity of the user obtained by applying an Extended Kalman Filter (EKF) (cf. Sect. IV) to the dynamic model of the human (1). Note that $\omega_h^*(t + \Delta t)$ is the angular velocity that the user should have at time $t + \Delta t$ in order to properly follow the robot. Let $\alpha, \delta \in \mathbb{R}^+$, the

proposed haptic feedback policy is,

$$\begin{aligned} s_L(t) &= \begin{cases} \text{on,} & \text{if } \omega_h^*(t + \Delta t) - \hat{\omega}_h(t + \Delta t) > \alpha \\ \text{off,} & \text{else,} \end{cases} \\ s_R(t) &= \begin{cases} \text{on,} & \text{if } \omega_h^*(t + \Delta t) - \hat{\omega}_h(t + \Delta t) < -\alpha \\ \text{off,} & \text{else,} \end{cases} \\ s_S(t) &= \begin{cases} \text{on,} & \text{if } \|\mathbf{P}_r(t) - \mathbf{P}_h(t)\| < \delta \\ \text{off,} & \text{else.} \end{cases} \end{aligned}$$

The threshold value α is used to avoid excessive alternation between the haptic stimuli *turn left* and *turn right*, as they can generate frequent and unwanted oscillations in the human locomotion. The signal $s_S(t)$ is sent to warn the human if she/he is too close to the robot, i.e., the actual human-robot distance is less than δ . In real scenarios the maximal robot velocities are limited. Thus, it may happen that the robot cannot maintain the formation, if the human moves too fast. The limit is: $v_r(t + \Delta t) > V_r$ (cf. (4)), where $V_r \in \mathbb{R}^+$ represents the maximal linear velocity of the robot. Note that the human user is always free to decide her/his pace. Only when the minimal human-robot distance is violated, a proper haptic signal is sent to the user in order to inform her/him to slow down.

Concerning the configuration with a single armband, the three stimuli (L , S , R) are mapped one-to-one onto the three tactors (left, center, right) of the device. The user wears the armband, as depicted in Fig. 4(a): the tactors representing the right and left direction are on the corresponding sides of the forearm. We assume that the orientation of user's arm does not vary too much during the motion, since it may influence the left/right location of the tactors. This issue could be solved by modifying the tracking algorithm in Sect. IV in order to estimate the orientation of the forearm. Thus, vibrotactile stimuli could be dynamically mapped on the three tactors depending on their actual positions. In the bilateral configuration, vibration of the left armband alerts the participant to turn left, and vice versa. The *slow down* stimulus is displayed by a vibration of both armbands.

In order to reduce the *aftereffect* problem (Pacinian corpuscles that sense vibration on the skin may adapt to continuous stimuli, see [30] and the references therein) and to preserve the users' ability to localize vibration, in both configurations we activate the tactors with a square wave. It has period 2τ , $\tau \in \mathbb{R}^+$, duty cycle of 50%, and logic levels 0 and 1 (Fig. 5(top)). When the logic level is high, the tactor vibrates with a frequency of 280 Hz (which is in the range of maximal sensitivity [28], [29]), and amplitude of 0.6 G (which is the maximal amplitude exerted by the tactors). On the contrary, when the logic level is low, the tactor is turned off. For the bilateral configuration, two tactors alternatively vibrate when a stimulus is sent to the device. In other words, the squared waves sent to the tactors are shifted by τ (Fig. 5). Note that in our application, the situation in which all tactors are turned on, or when the left and right tactors are simultaneously activated, never occurs.

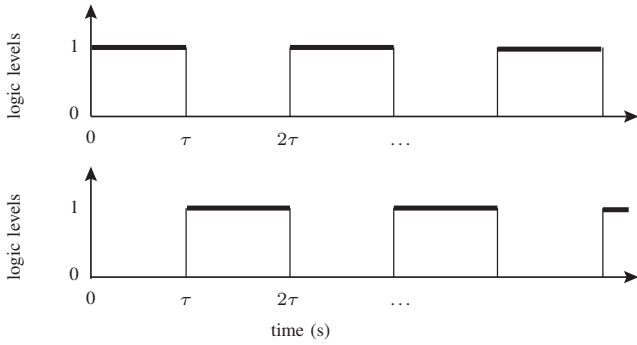


Fig. 5. *Vibrotactile stimuli*. In order to preserve the users' ability to localize vibrations, in all configurations each factor is activated with a square wave having period 2τ and duty cycle of 50% (top). When the logic level is 1, the tactor vibrates with a frequency of 280 Hz and amplitude of 0.6 G. When the logic level is 0, the tactor is turned off. If two tactors are simultaneously activated, their signals are shifted by τ (bottom).

C. Evaluation of the haptic feedback

The proposed device was tested on 7 healthy subjects (6 males, age range 23-40, 5 right-handed). 2 of them had experience with previous prototypes of our haptic armband (based, however, on different electromechanical designs). None of the participants reported any deficiencies in the perception abilities (including vision, hearing, touch and proprioception). The participants signed an informed consent form. All of them were informed about the purpose of the experiment. They were allowed to discontinue participation at any time. No payment was provided for the participation.

Two different experiments were performed. The aim of the first one was to evaluate whenever the unilateral and bilateral conditions could elicit the intended causal chain of stimulus-perception-response. The second experiment was performed to evaluate the maximal stimulus duration that did not degrade the perception of the stimulus itself, since vibration effects may persist after the end of the stimulation (*aftereffect* problem). In order to evaluate the users' experience, a questionnaire using bipolar Likert-type five-point scales was filled out by each subject at the end of the experiments for both haptic conditions.

In the first experiment, participants were instructed to walk along a walkway whilst wearing the armband/s, and to react accordingly to the stimulus type (L , S , R), as soon as they perceived it. The length of the walkway was about 4 m. The vibrotactile stimulus was provided as soon as the user was 1.7 m in front of the obstacle. The armband/s continued to vibrate for 2 s. For each haptic configuration (unilateral and bilateral), each subject performed 12 trials (4 trials for each stimulation type) organized in a pseudo-random order. All subjects were blindfolded and wore circumaural headphones reproducing white noise to mask the distracting ambient or cueing sounds from the stimulators. Two RGB-D cameras tracked the motion of the human by using a custom designed tracking algorithm (Sect. IV). Sequences of stimulation appeared in short bursts with $\tau = 0.2$ s, vibration frequency of 280 Hz, and amplitude of 0.6 G (Fig. 5). The vibration period 2τ was determined both by mechanical limitation of the proposed tactors, and by pilot studies. Such experiment allowed to evaluate the haptic devices in a scenario as similar as possible to the final setup.

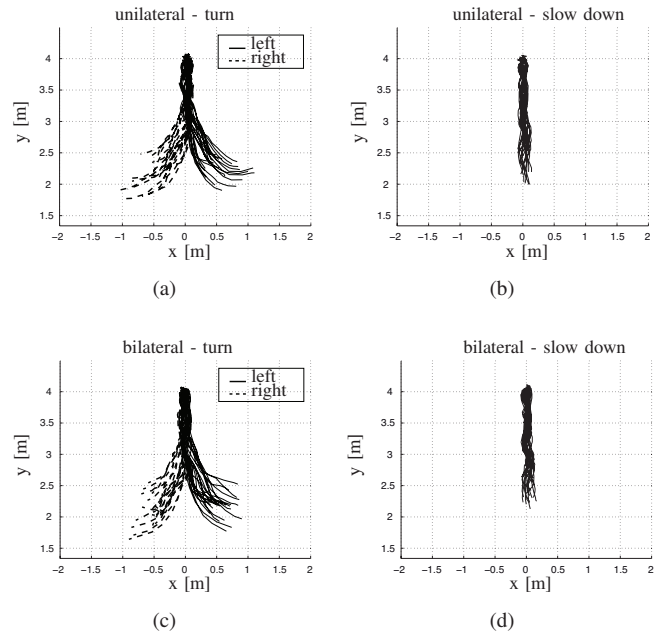


Fig. 6. *Evaluation of the haptic feedback*. Trajectories performed by the users, as the participants walk from top to bottom using a single vibrotactile armband (a)-(b) and two armbands (c)-(d), for the three stimuli (turn left, turn right, and slow down), respectively.

In the second experiment, we analyzed if a stimulus with a long duration affected the perception of the stimulus itself (*aftereffect* problem). Each subject was seated comfortably at a desk. Both feedback conditions (unilateral and bilateral) were evaluated. The subjects were given circumaural headphones emitting white noise to mask distracting sounds. Each subject tested two sets of vibrotactile stimuli. Each set was composed of pulsed square wave signals with period $2\tau = 0.4$ s, amplitude of 0.6 G (Fig. 5), and 4 different durations (2 s, 10 s, 30 s and 60 s) unknown to the users. Each signal was displayed to the armband. The user had to notify when the armband stopped to vibrate. For each vibrotactile stimulus, we recorded the interval of time between the end of the stimulus and the instant in which the user notified it. Responses were made by pressing a specific button on a keypad. For each stimulus, we asked the users if they felt any tingling sensation.

The questionnaire, consisting of 6 questions, was designed to evaluate their comfort, opinion of feedback quality, perceived effectiveness of the feedback, intrusiveness and flexibility of the device, and overall preferences. An answer of 5 meant *strongly agree*, whereas an answer of 1 meant *strongly disagree*.

1) *Data analysis*: In the first experiment, all subjects correctly reacted to the proposed stimuli for both haptic configurations (Fig. 6). By analyzing the trajectories performed by the participants, we estimated their reaction time (time

TABLE I
AVERAGE REACTION TIME OF THE USERS FOR THE GIVEN STIMULI (TURN LEFT, TURN RIGHT, SLOW DOWN).

Condition	Turn left (s)	Turn right (s)	Slow down (s)
Unilateral	0.86 ± 0.13	0.80 ± 0.14	0.87 ± 0.16
Bilateral	0.72 ± 0.14	0.74 ± 0.16	0.75 ± 0.16

TABLE II
AVERAGE INTERVAL OF TIME BETWEEN THE END OF THE STIMULUS
(DURATION 2 s, 10 s, 30 s, 60 s) AND THE INSTANT IN WHICH THE USER
NOTIFIED IT.

Condition	2 s (s)	10 s (s)	30 s (s)	60 s (s)
Unilateral	0.72 ± 0.08	0.70 ± 0.19	0.67 ± 0.19	0.78 ± 0.12
Bilateral	0.72 ± 0.14	0.75 ± 0.14	0.69 ± 0.24	0.79 ± 0.14

taken for the users to make a turn, or to slow down, after the stimulus was sent) (Table I). The average reaction time was approximately 0.84 s with a standard deviation of 0.22 s using the single armband and 0.74 s with a standard deviation of 0.21 s using two armbands. Comparison of the means among the feedback conditions was tested using a two-way repeated-measures ANOVA (ANalysis Of VARIance) [31]. Feedback conditions and localization of the feedback signals (L , S , R) were considered as within-subject factors. A family-wise level $\alpha_p = 0.05$ was used for all tests. The collected data passed the Shapiro-Wilk normality test and the Mauchly's Test of Sphericity. The means did not differed significantly among feedback conditions [$F(1, 6) = 3.905$, $p = 0.096$, $\alpha_p = 0.05$], meaning that the reaction time of the users was not influenced by using the unilateral or bilateral condition to present directional cues. For each feedback condition, a one-way repeated measures ANOVA was conducted to determine whether reaction times for different stimuli (L , S , R) changed in a significant way. In both conditions, the collected data passed the Shapiro-Wilk normality test and the Mauchly's Test of Sphericity. Tests showed that reaction times for the given stimuli did not depend on the type of stimulus: unilateral condition [$F(2, 12) = 1.853$, $p = .199$, $\alpha_p = 0.05$], bilateral condition [$F(2, 12) = 0.154$, $p = .859$, $\alpha_p = 0.05$].

The authors are aware that the proposed tests were conducted in a controlled environment and performed on healthy, able bodied adults and that the reaction time may increase in a real world situation, such as a loud factory or a busy hospital and with older/impaired subjects. Nevertheless, tests performed in Sect. V show the validity of our approach in a real scenario.

In the second experiment, we performed a one-way repeated-measures ANOVA on the elapsed time to determine whether reaction times for different stimulus durations (2 s, 10 s, 30 s, 60 s) were related to the duration of the stimulus itself (Table II). For both feedback conditions, the collected data passed the Shapiro-Wilk normality test and the Mauchly's Test of Sphericity. Tests showed that reaction times did not depend on the duration of the stimulus: unilateral condition [$F(3, 18) = 0.421$, $p = 0.740$, $\alpha_p = 0.05$], bilateral condition [$F(2, 12) = 1.853$, $p = .199$, $\alpha_p = 0.05$]. Since no tingling sensation was felt by the users, we can state that in our application the aftereffect problem never occurs as long as signals with duration lower than 60 s are considered.

2) *Survey responses*: A questionnaire, presented in the form of bipolar Likert-type five-point scales (Table III), was filled out by the users in order to understand how they judged the two different feedback configurations. First four questions U1-4 and B1-4 investigated how much the users found the two configurations usable and comfortable. Questions U5-6 and

TABLE III
QUESTIONNAIRE PROPOSED AT THE END OF THE EXPERIMENTS FOR THE
UNILATERAL AND BILATERAL CONDITION, RESPECTIVELY.

Questions	
U1	The unilateral condition is easy to use.
U2	The unilateral condition is not hampering.
U3	Following the cues of the unilateral condition is not tiring.
U4	Wearing one single armband is a comfortable solution.
U5	The cues suggested by the unilateral condition give comprehensive information for the guidance system.
U6	The cues suggested by the unilateral condition are easy to distinguish.
B1	The bilateral condition is easy to use.
B2	The bilateral condition is not hampering.
B3	Following the cues of the bilateral condition is not tiring.
B4	Wearing two armbands is a comfortable solution.
B5	The cues suggested by the bilateral condition give comprehensive information for the guidance system.
B6	The cues suggested by the bilateral condition are easy to distinguish.

TABLE IV
RESULTS OF THE QUESTIONNAIRE FOR THE UNILATERAL AND BILATERAL
CONDITION, RESPECTIVELY.

Questions	Results	Questions	Results
U1	3.86 ± 0.690	B1	4.00 ± 0.816
U2	4.71 ± 0.488	B2	4.29 ± 0.488
U3	4.29 ± 0.756	B3	4.14 ± 0.690
U4	4.14 ± 0.690	B4	2.86 ± 0.690
U5	3.86 ± 0.690	B5	3.71 ± 0.756
U6	2.86 ± 0.690	B6	4.29 ± 0.756

B5-6 investigated if the users felt the suggested cue to be informative enough and if the cues were easy to distinguish in the two configurations. A series of Wilcoxon Signed-Rank tests was performed for highlighting statistical significance of the difference between the proposed questions (Table IV). No significant differences were found between question U1-3 and B1-3 and between question U5 and question B5, showing that the two configurations were easy to use, not tiring and did not hamper the user. Moreover, the haptic cues sent to the participants were found informative enough. Eventually, the unilateral solution was considered comfortable [$Z = -2.251$, $p = 0.024$, $\alpha_p = 0.05$], whereas the cues sent through the bilateral solution were easier to understand and more intuitive [$Z = -2.060$, $p = 0.039$, $\alpha_p = 0.05$].

The proposed feedback configurations were comparable (see also Sect. III-C1). In the experimental validation of the system, we decided to use the bilateral configuration, since cues sent through this solution were found easier to understand. We believe that the results obtained in the experimental validation would have not differed too much if the unilateral condition was used.

IV. VISUAL DETECTION AND TRACKING OF THE HUMAN

The human-robot control policy described in (4)-(5) requires an estimation of the human's pose and velocities. This section provides an overview of the major steps of our method for estimating such parameters from dense depth images provided by an RGB-D camera on-board the robot. We believe that our approach is relatively general and can also be applied to other typologies of vision sensors (e.g., time-of-flight cameras).

A. Description of the tracking algorithm

Let $\tilde{\mathbf{x}}_h = [\tilde{x}_h, \tilde{y}_h, \tilde{\theta}_h, \tilde{v}_h, \tilde{\omega}_h]^T$ be the state of the human that we need to estimate. The first step of the tracking

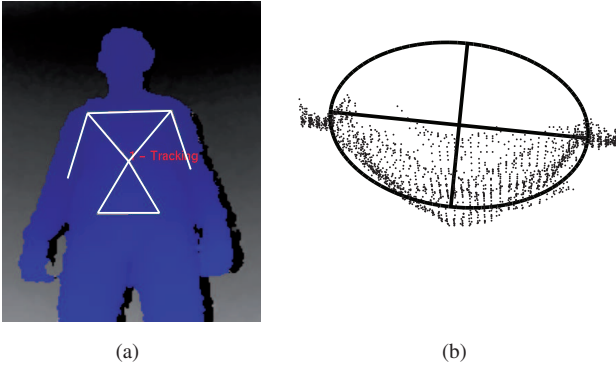


Fig. 7. *Human body tracking method on real data.* (a) NITE's skeleton tracker is used to initially detect the subject (the skeleton of the torso is shown in white). (b) Data points which are too far from the torso are removed, while the remaining points are down-sampled and expressed in the robot reference frame. The points are finally projected onto the robot's plane of motion and the pose of the human is detected via ellipse fitting.

algorithm consists in determining the pose of the user. We use NITE skeleton tracker to initially detect the human (Fig. 7(a)) and Point Cloud Library [32] to process the depth (3-D point) data and extract the information of the human pose. The shoulders play an important role in the description of the human locomotion (cf. [18]). This is why in the detection phase we discard all the 3-D points that are too far from the human's torso. In order to speed up the tracking process, we first down-sample the data using a voxel grid filter with a leaf size of 1 cm. Then, we express the down-sampled point cloud in the robot reference frame, and we project the point cloud onto the robot's plane of motion (xy -plane). Finally, an ellipse fitting [33] is performed over the projected points (Fig. 7(b)). The estimated position $(\tilde{x}_h, \tilde{y}_h)$ of the human corresponds to the center of the ellipse. The user's orientation $\hat{\theta}_h$ is assumed to be coincident with the orientation of the major axis of the ellipse with respect to the robot's frame. In order to fully exploit the temporal information inherent to human's motion and to estimate $\tilde{\mathbf{x}}_h$ from pose measurements, we implement an Extended Kalman Filter (EKF). EKF provides an estimation of the current state $\tilde{\mathbf{x}}_h$ as well as a one-step ahead prediction of it. The prediction of $\tilde{\mathbf{x}}_h$ is used to generate suitable haptic signals (cf. Sect. III-B).

In case of failures of the skeleton tracker, we select the 3-D points in the neighborhood of the predicted human pose. We project such points onto the robot's plane of motion. Successively, we perform a cluster filtering to discard those clusters whose dimension is outside of a given range, and whose distance is far enough from the last tracked human position. Finally, an ellipse fitting is performed over the resulting cluster. An example of this procedure is visible in the attached video.

B. Evaluation of the tracking algorithm

The proposed method runs at an average frame rate of 27 frames per second on a laptop with 16 GB RAM, 2.4 GHz Intel i7 CPU, and NVIDIA GeForce GTX 765M graphic card. Synthetic data with ground truth information were used for the quantitative evaluation of the proposed method. This is a common approach in the relevant literature because ground

truth data for real-world image sequences is hard to obtain. The employed synthetic sequence consisted of 7 trajectories, each one composed of 60 consecutive human poses (a total of 420 poses were considered) that encoded the human walking motion. All the trajectories lied in a $3 \text{ m} \times 3 \text{ m}$ area. The user's heading ranged from -90 deg to 90 deg . Computer graphic was used to synthesize the required input for each considered pose. The method was also evaluated with respect to its tolerance to noisy observations. Two types of noise were considered: errors in depth estimation and errors in the camera orientation with respect to the floor. The latter ones affected the correct projection of the point cloud onto the robot xy -plane. We considered the camera orientation error as noise on the roll angle of the camera frame. We modeled the errors as Gaussian distributions centered around the actual value with the variance controlling the amount of noise.

Figs. 8 (a)-(b) show the mean and the standard deviation of both the pose-estimation error $\|(x, y)^T - (\tilde{x}, \tilde{y})^T\|$ and the heading-estimation error $|\theta - \hat{\theta}|$, when noise was added to depth estimation. Figs. 8 (c)-(d) show the mean and the standard deviation of pose and heading-estimation error, when noise was added to the estimation of the floor orientation with respect to the camera frame. From Fig. 8 we observe that the performance of our tracker is not critically affected by errors in depth estimation, or in camera roll angle estimation.

V. EXPERIMENTAL VALIDATION

We tested the proposed control strategy (4) in an indoor environment using a Pioneer LX robot (with maximal linear velocity of 1.8 m/s) equipped with a backward facing Asus Xtion RGB-D camera (Fig. 9).

Fifteen healthy subjects (age range 23-52, 12 males, 13 right-handed) were involved in our experiments¹. Five of them participated in the evaluation of the haptic armband (cf. Sect. III-C). None of the participants reported any deficiencies in the perception abilities (including vision, hearing, touch and proprioception). The participants signed an informed consent form. All of them were informed about the purpose of the experiment, were able to discontinue participation at any time and no payment was provided for the participation. All subjects were blindfolded and instructed to move accordingly to the haptic feedback, but no instructions were given about their velocities. Since the surrounding sounds could probably modify the users' behavior, as they could be afraid to hit something, we cut off the auditory feedback by reproducing white noise through earphones.

Two different trajectories (clockwise and counterclockwise) were considered for the robot in each modality (Fig. 10). The trajectories were about 225 m and 223 m long, respectively. Each one was composed of four clear long corridors (the width of the corridors ranged from 1.2 to 2.2 m) and six 90 deg turns. Each subject performed 4 trials, 2 for each trajectory, in a randomized order. Thus the total number of considered trials was 60. In order to evaluate the proposed haptic policy, the subjects additionally performed 2 trials for each trajectory. In

¹Please notice that this paper is accompanied by multimedia material. The videos of the real-time experiments are available also at: <http://goo.gl/aH1yO2>

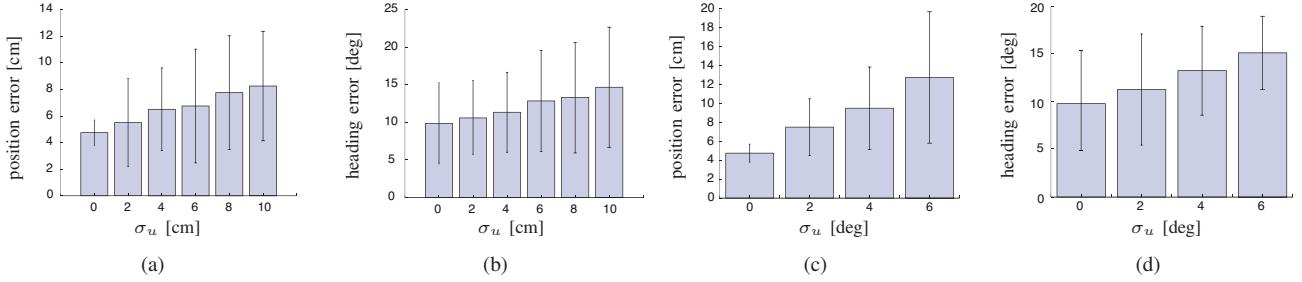


Fig. 8. Evaluation of the tracking algorithm. Mean and standard deviation of the position and heading estimation error with increasing noise on: (a)-(b) depth estimation; (c)-(d) camera roll angle estimation. The noise was modeled as a Gaussian distribution centered in the actual value with variance σ_u . The supplemental material accompanying the paper provides videos with the results obtained in real-world sequences.

this case, the users had no vision impairment, and the desired trajectory was displayed by the laptop positioned on the mobile robot. It is worth noting that the experiments with blindfolded people were performed to show the validity of the proposed approach in the challenging scenario in which visual and also auditory information might not be available. In other terms, blind-folding was meant more to prove how performant was our method more than specifically investigating guidance for blinds.

The robot had a map of the environment and autonomously localized itself via the Monte Carlo Localization [34] provided by the Aria Core Library [35]. The initial obstacle-free paths for both the robot and the user were computed offline using a customized version of the planner presented in [36]. In addition, we considered 3 static virtual obstacles and 2 dynamic ones (Figs. 10(a)-(b)). The obstacles were unknown to the robot, i.e., the initial paths did not consider such obstacles. We simulated a sensing range of 4 m for the robot. As soon as the obstacles were within the sensing range of the robot, the actual path was updated online by running a new instance of the planner. The camera was rotated about its x -axis of 23.20 deg. The formation parameters were set to $l^d = 1.1$ m and $\psi^d = \pi$, $k_1 = k_2 = 3$, $d = 0.1$ m, $\alpha = 0.7$ rad/s and $\delta = l^d - 0.2$ m. The parameters above were determined by both the mechanical limitations of the system and the environment. They were set in order to allow the user to properly navigate and accomplish the goal.

For each trajectory we computed the formation error $\mathbf{E}(t) = \mathbf{P}_h(t) - \mathbf{P}_r(t) - l^d (\cos \psi^d, \sin \psi^d)^T$. Fig. 11 shows the trials in which lowest formation error was achieved. Figs. 11(a)-(d) report the actual position of the reference point $\mathbf{P}_h(t)$ and its desired pose computed as $\mathbf{P}_r(t) + l^d (\cos \psi^d, \sin \psi^d)^T$. Figs. 11(b)-(e) show the time evolution of the norm of the formation error $\mathbf{E}(t)$ for both trajectories. Peaks in the formation error were mainly due to the rotational velocity of the robot in correspondence of sharp turns, and to inaccurate estimations of human's pose. The related vibrational signals of the haptic devices are reported in Figs. 11(c)-(f). Figs. 12(a)-(d) show the formation error for each trial $\mathbf{E}^i(t)$, $i = 1, 2, \dots, 60$. The percentage of the total duration of the trial in which the vibrotactile armband was activated is reported in Figs. 12(b)-(e). Finally Figs. 12(c)-(f) report the mean (and the standard deviation) of the linear velocity of the users for each trial $v_h^i(t)$, $i = 1, 2, \dots, 60$. In white are reported the results of the subjects who had previously participated in the evaluation

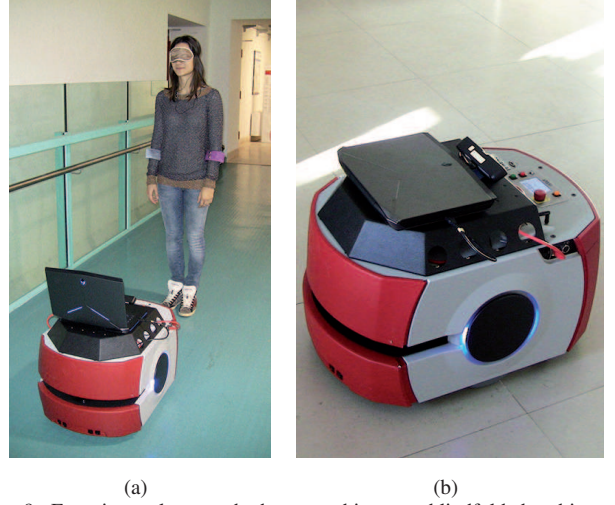


Fig. 9. Experimental setup: the human subject was blindfolded and instructed to move accordingly to the haptic feedback provided by two custom-design vibrotactile armbands. (b) Pioneer LX robot equipped with a backward facing Asus Xtion camera.

of the haptic devices (cf. Sect. III-C). For both trajectories the mean of the formation error is always smaller than 0.3 m. Moreover, users who never tried the haptic interface before were able to correctly recognize the haptic stimuli and follow the robot. Results show the functionality of the proposed approach. For the clockwise and counterclockwise trajectories, the mean of the formation error $\mathbf{E}^i(t)$ among all the trials was 0.24 ± 0.04 m and 0.23 ± 0.05 m. The average percentage of time in which the armbands were turned on was 26.65 ± 7.10 % and 24.41 ± 6.91 %, while the average of the users' linear velocities was 0.62 ± 0.07 m/s and 0.63 ± 0.08 m/s. Concerning the activation time of the armbands it is worth noting that also during a straight line, the armbands may correct the trajectory of the users due to the well known fact that it is hard for a blindfolded people to walk exactly straight, due to the absence of landmarks. Thus, also a straight line can reveal if the proposed approach is valid. For the linear velocities of the subjects, it is worth noting that we asked the subjects to walk at their comfortable speed. Moreover, due to the reduced activation of the *slow down* behavior, the users' linear velocities were mainly determined by the confidence of the users in the system. Experiments performed on users with no vision impairments revealed that for the clockwise and counterclockwise trajectories the formation error $\mathbf{E}^i(t)$ among all the trials was 0.15 ± 0.03 m and

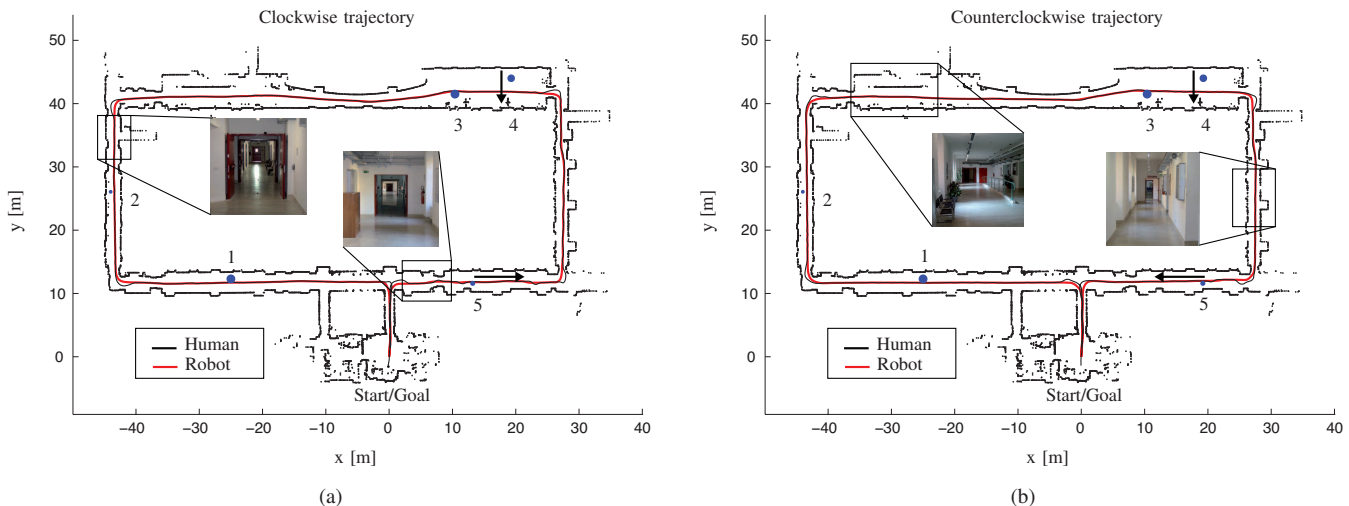


Fig. 10. *Experimental scenario.* The experimental scenarios contained a goal region and virtual dynamic and static obstacles (blue). The objective was to guide the user from her/his starting position to the goal one while avoiding obstacles. We considered three static obstacles (obstacles 1, 2, and 3) and two moving obstacles (obstacles 4 and 5). The black arrows represent the velocity directions of the dynamic obstacles. The speed of the moving obstacles was 0.4 and 0.5 m/s, respectively for obstacle 4 and 5. Each user performed the proposed trajectory four times: two times in a clockwise order (a) and twice in a counterclockwise order (b). A path planner was used to generate the initial trajectories for both the robot and the user. The initial trajectories did not consider the obstacles which were unknown to the users. For the obstacles, we considered a sensing range of 4 m for the robot, i.e., when an obstacle was inside the sensing range, the planner was used to update the current trajectories. The insets show sample images of the environment.

0.13 ± 0.02 m, and the average of the users' linear velocities was 0.82 ± 0.08 m/s and 0.79 ± 0.07 m/s. A Paired Samples T-Test revealed no statistical differences between the clockwise and counterclockwise trajectories (for both the vision-impaired and no impairment condition) in terms of formation error, users' walking speeds, and activation time of the armbands. Smaller formation errors (clockwise trajectory $t_{14} = -9.225$, counterclockwise trajectory $t_{14} = -9.187$) and faster paces (clockwise trajectory $t_{14} = 15.663$, counterclockwise trajectory $t_{14} = 8.198$) were found for the no impairment condition with respect to the vision-impaired one, $p < 0.005$.

A. Discussion

Although the results presented in Sect. V are promising, a comparison between the results obtained using this approach and experiments performed with sighted people reveal that additional studies need to be done in order to have comparable formation errors and walking speeds.

The proposed haptic feedback policy assumes that users behave like unicycle systems and smoothly rotate when a proper vibrotactile stimuli is received. Under these assumptions, the haptic feedback can direct the user toward the desired pose until she/he is close enough to it. If the user sharply turns when a stimuli is received, it may become difficult for the proposed method to correctly guide her/him, mainly due to delays in the reaction time of the user. However, this situation never happened in our experimental validation.

We designed the system in a way that the user always remains in charge of the final decision to take and she/he can always override the *suggestions* given by the system. A possible drawback of such decision is that, in case of danger, the proposed system can not force the user to move in a particular way. This problem is indeed shared among all the approaches that use tactile feedback.

The Asus Xtion offers a 58 deg horizontal wide viewing angle and an effective sensing range of 0.8-3.5 m. It works

well in an almost completely open environment, however its real world uses can be limited. In the proposed experiments, we showed that it is possible to use such sensor also in less open environments. The human was correctly tracked around a series of 90 deg turns through hallways by using a proper choice of formation parameters and trajectory for the robot. It is worth noting that the formation parameters (l^d , ψ^d) should be accurately tailored depending on the sensors' characteristics and on the environment.

VI. CONCLUSION AND FUTURE WORK

The paper presents a new paradigm for the assisted navigation of a human in an unknown environment with a mobile robot using visual and haptic information. The subject is free to decide her/his own pace. A warning vibrational signal is generated by haptic armbands only when a large deviation with respect to the planned route occurs. In this work we show that, based on recent studies, control algorithms designed for robot teams can be suitable applied to human-robot formations. Moreover, we demonstrate that by exploiting the nonholonomic nature of human locomotion, few vibrotactile stimuli are sufficient to effectively guide the user in mixed human-robot formations. Finally, our cooperative guidance system is easy to use and it does not need long training programs. The effectiveness of the proposed approach is demonstrated via real-world experiments conducted on 15 subjects in a large indoor environment.

The use of wearable haptic devices opens new scenarios in mixed human-robot teams. In future work, we plan to improve the haptic feedback in order to make the users more confident about the system. We will investigate the extension of the proposed work to a mixed team composed of a robot and multiple users. Similarly, we will extend the proposed haptic feedback to more complex tasks like cooperative manipulation between a human and a mobile manipulator.

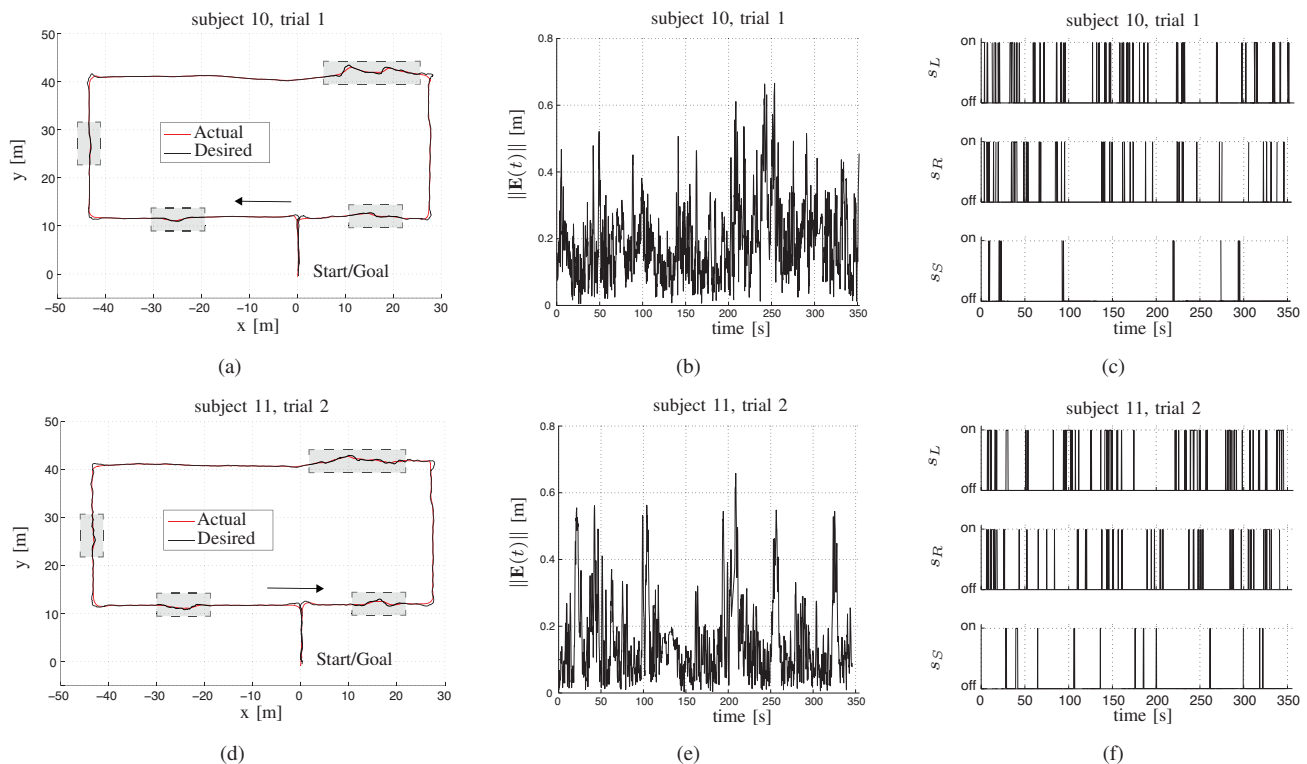


Fig. 11. *Experimental results.* Clockwise (top) and counterclockwise trajectories (bottom). (a)-(d) Desired and actual trajectories performed by the users: the shaded areas represent the portions of the trajectory which were updated due to the presence of static and dynamic obstacles; (b)-(e) formation error $\mathbf{E}(t) = (E_x(t), E_y(t))^T$; (c)-(f) armbands activation time for the users who achieved the lowest formation error.

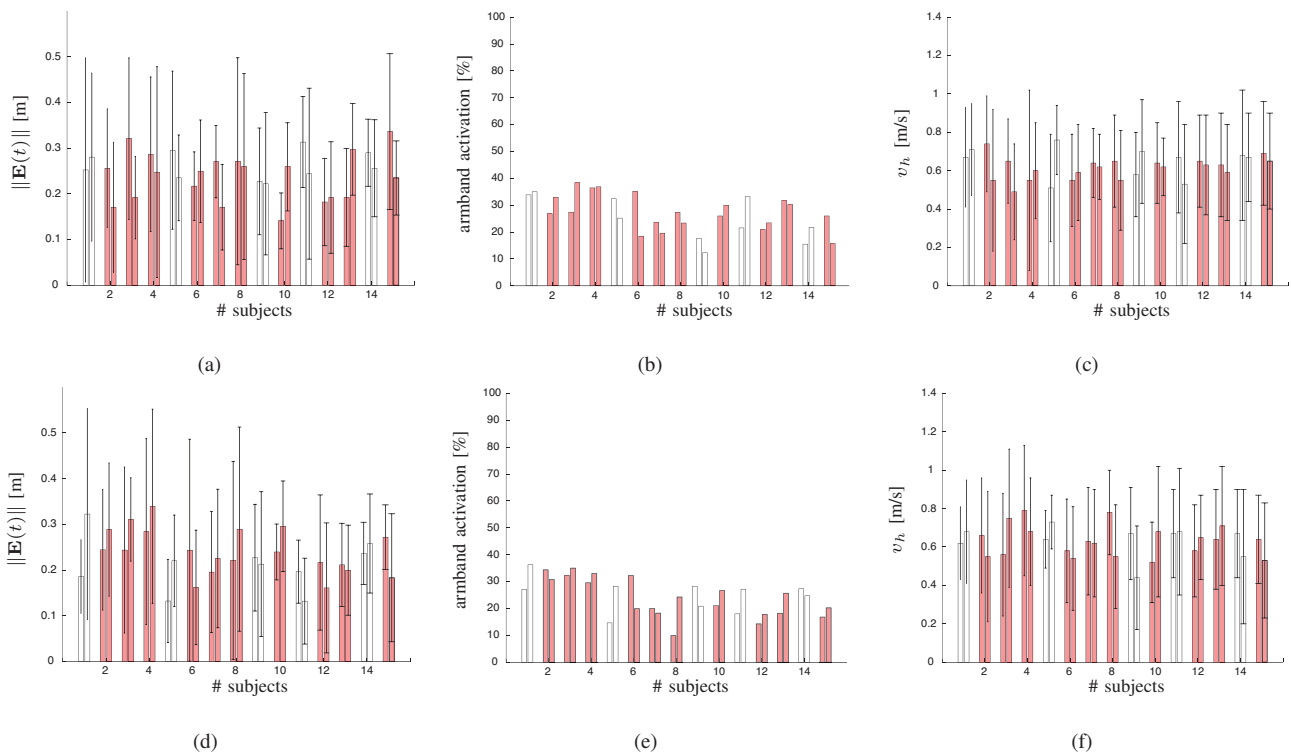


Fig. 12. *Experimental results.* Clockwise (top) and counterclockwise (bottom) trajectories. (a)-(d) Mean and standard deviation of the norm of the formation error $\mathbf{E}(t) = (E_x(t), E_y(t))^T$ over the 60 trials for the 15 subjects. (b)-(e) Percentage of activation time of the armbands with respect to the trajectory execution time. (c)-(f) Mean and standard deviation of the linear velocity $v_h(t)$ of the users. In white are reported the subjects who participated in the evaluation of the haptic armband.

REFERENCES

- [1] J. Lieberman and C. Breazeal. TIKL: Development of a wearable vibrotactile feedback suit for improved human motor learning. *IEEE Trans. Robot.*, 23(5):919–926, 2007.
- [2] M. A. Heller and W. Schiff. *The psychology of touch*. Psychology Press,

- 2013.
- [3] D. Fontanelli, A. Giannitrapani, L. Palopoli, and D. Prattichizzo. Unicycle steering by brakes: a passive guidance support for an assistive cart. In *Proc. 52nd IEEE Conf. Dec. Contr.*, pages 2275–2280, 2013.
- [4] M. Peshkin, J. Colgate, W. Wannasuphprasit, C. Moore, R. Gillespie, and P. Akella. Cobot architecture. *IEEE Trans. Robot. Autom.*, 17(4):377–390, 2001.
- [5] A.K. Das, R. Fierro, V. Kumar, J.P. Ostrowski, J. Spletzer, and C.J. Taylor. A vision-based formation control framework. *IEEE Trans. Robot. Autom.*, 18(5):813–825, 2002.
- [6] A. Ghosh, L. Alboul, J. Penders, P. Jones, and H. Reed. Following a robot using a haptic interface without visual feedback. In *Proc. Int. Conf. on Advances in Computer-Human Interactions*, pages 147–153, 2014.
- [7] G. Galatas, C. McMurrough, G.L. Mariottini, and F. Makedon. eyeDog: an assistive-guide robot for the visually impaired. In *Proc. 4th Int. Conf. Perv. Tech. Assist. Env.*, pages 58:1–58:8, 2011.
- [8] A. Ranasinghe, J. Penders, P. Dasgupta, K. Althoefer, and T. Nanayakkara. A two party haptic guidance controller via a hard rein. In *Proc. IEEE/RSJ Int. Conf. Intel. Robots Syst.*, pages 116–122, 2013.
- [9] A. Ranasinghe, P. Dasgupta, K. Althoefer, and T. Nanayakkara. Identification of haptic based guiding using hard reins. *PLoS one*, 10(7), 2015.
- [10] R.W. Lindeman, Y. Yanagida, H. Noma, and K. Hosaka. Wearable vibrotactile systems for virtual contact and information display. In *Virtual Reality*, volume 9, pages 203–213. Springer, 2006.
- [11] A.P. Arias and U.D. Hanebeck. Wide-area haptic guidance: Taking the user by the hand. In *Proc. IEEE/RSJ Int. Conf. Intel. Robots Syst.*, pages 5824–5829, 2010.
- [12] J. B. F. Van Erp, H. A. H. C. Van Veen, Chris C. Jansen, and T. Dobbins. Waypoint navigation with a vibrotactile waist belt. *ACM Trans. Appl. Percept.*, 2(2):106–117, 2005.
- [13] A. Cosgun, E.A. Sisbot, and H.I. Christensen. Guidance for human navigation using a vibro-tactile belt interface and robot-like motion planning. In *Proc. IEEE Int. Conf. Robot. Autom.*, pages 6350–6355, 2014.
- [14] R. L. Koslover, B. T. Gleeson, J. T. de Bever, and W. R. Provancher. Mobile navigation using haptic, audio, and visual direction cues with a handheld test platform. *IEEE Trans. Haptics*, 5(1):33–38, 2012.
- [15] J. Dumora, F. Geffard, C. Bidard, T. Brouillet, and P. Fraisse. Experimental study on haptic communication of a human in a shared human-robot collaborative task. In *Proc. IEEE/RSJ Int. Conf. Intel. Robots Syst.*, pages 5137–5144, 2012.
- [16] D. Sieber, S. Music, and S. Hirche. Multi-robot manipulation controlled by a human with haptic feedback. In *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, Hamburg, Germany, 2015.
- [17] J. Stückler and S. Behnke. Following human guidance to cooperatively carry a large object. In *Proc. 11st IEEE-RAS Int. Conf. Humanoid Robots*, pages 218–223, 2011.
- [18] G. Arachavaleta, J.P. Laumond, H. Hicheur, and A. Berthoz. On the nonholonomic nature of human locomotion. *Auton. Robot.*, 25(1):25–35, 2008.
- [19] S. Scheggi, F. Chinello, and D. Prattichizzo. Vibrotactile haptic feedback for human-robot interaction in leader-follower tasks. In *ACM PETRA Workshop on Robotics in Assistive Environments*, Article No. 51, 2012.
- [20] S. Scheggi, M. Aggravi, F. Morbidi, and D. Prattichizzo. Cooperative human-robot haptic navigation. In *Proc. IEEE Int. Conf. on Robotics and Automation*, pages 2693–2698, 2014.
- [21] J.-J.E. Slotine and W. Li. *Applied nonlinear control*, volume 199. Prentice-Hall Englewood Cliffs, NJ, 1991.
- [22] C. Canudas de Wit, H. Khenouf, C. Samson, and O. J. Sordalen. *Nonlinear control design for mobile robots*. World Scientific Series in Robotics and Automated Systems, Vol. 11. World Scientific, 1993.
- [23] F. Gemperle, T. Hirsch, A. Goode, J. Pearce, D. Siewiorek, and A. Smailigic. Wearable vibro-tactile display, 2003. Carnegie Mellon University.
- [24] I. Karuei, K.E. MacLean, Z. Foley-Fisher, R. MacKenzie, S. Koch, and M. El-Zohairy. Detecting vibrations across the body in mobile contexts. In *Proc. Int. Conf. Human Factors Comp. Syst.*, pages 3267–3276, 2011.
- [25] L. J. Post, I. C. Zompa, and C. E. Chapman. Perception of vibrotactile stimuli during motor activity in human subjects. *Exp. Brain Res.*, 100(1):107–120, 1994.
- [26] R. W. Cholewiak and A. A. Collins. Vibrotactile localization on the arm: Effects of place, space, and age. *Perception & Psychophysics*, 65(7):1058–1077, 2003.
- [27] S. Weinstein. Intensive and extensive aspects of tactile sensitivity as a function of body part, sex, and laterality. In Spriengfield, editor, *The skin senses*, pages 195–218. Erlbaum, 1968.
- [28] A. Riener. Sensor-actuator supported implicit interaction in driver assistance systems. In S. Hölldobler et al., editor, *Ausgezeichnete Informatikdissertationen 2009*, volume 10, pages 221–230. Gesellschaft für Informatik, Bonn, 2010.
- [29] R. W. Cholewiak and A. A. Collins. Sensory and Physiological Bases of Touch. In Morton A. Heller and William Schiff, editors, *The Psychology of Touch*, chapter 2, pages 23–60. Erlbaum, 1991.
- [30] I.M. Vogels, A.M. Kappers, and J.J. Koenderink. Haptic aftereffect of curved surfaces. *Perception*, 25(1):109–119, 1996.
- [31] R. G. Miller. *Beyond ANOVA: basics of applied statistics*. Chapman & Hall, 1997.
- [32] R. B. Rusu and S. Cousins. 3D is here: Point Cloud Library (PCL). In *Proc. IEEE Int. Conf. Robot. Automat.*, pages 1–4, 2011.
- [33] A. Fitzgibbon, M. Pilu, and R. B. Fisher. Direct least square fitting of ellipses. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 21(5):476–480, 1999.
- [34] S. Thrun, W. Burgard, and D. Fox. *Probabilistic robotics*. MIT press, 2005.
- [35] Aria Core Library. [Online]: <http://www.mobilerobots.com/>.
- [36] S. M. Lavalle. Rapidly-exploring random trees: A new tool for path planning. Technical report, Computer Science Dept., Iowa State University, 1998.



Stefano Scheggi (S'09 - M'12) received the M.S. and Ph.D. degrees in Computer Engineering from the University of Siena, Italy, in 2007 and 2012, respectively. In 2011, he was a Visiting Ph.D. Student at the Department of Computer Science, George Mason University, Fairfax, USA, under the supervision of Prof. Jana Košecá. Since 2012, he holds a postdoctoral position at the University of Siena, Italy. His research interests include computer vision, visual servoing, mobile robotics, haptics and augmented/virtual reality. He is a member of the

IEEE.



Marco Aggravi (S'11 - M'16) received his M.Eng. Degree (2011) on Computer Engineering and his Ph.D. in Information Engineering and Science (2016) from the University of Siena. He was a Visiting Ph.D. student at DLR Institute of Robotics and Mechatronics, Oberpfaffenhofen, Germany in 2011. He is currently a research assistant at the SIRSLab, University of Siena. His fields of research are human-robot interaction and haptic guidance methodologies. He is a member of the IEEE.



Domenico Prattichizzo (S'93 - M'95 - SM'15 - F'16) received the M.S. degree in Electronics Engineering and the Ph.D. degree in Robotics and Automation from the University of Pisa in 1991 and 1995, respectively. Since 2002 Associate Professor of Robotics at the University of Siena. Since 2009 Scientific Consultant at Istituto Italiano di Tecnologia, Italy. In 1994, Visiting Scientist at the MIT AI Lab. Guest Co-Editor of Special Issue "Robotics and Neuroscience" of the Brain Research Bulletin (2008). Co-author of the "Grasping" chapter of "Handbook of Robotics" Springer, 2008, awarded with two PROSE Awards presented by the American Association of Publishers. Since 2014, Associate Editor of Frontiers of Biomedical Robotics. From 2007 to 2013 Associate Editor in Chief of the IEEE Trans. on Haptics. From 2003 to 2007, Associate Editor of the IEEE Trans on Robotics and IEEE Trans. on Control Systems Technologies. From 2013 Chair of the IEEE RAS Early Career Awards Evaluation Panel. Vice-chair for Special Issues of the IEEE Technical Committee on Haptics (2006-2010). Chair of the Italian Chapter of the IEEE RAS (2006-2010), awarded with the IEEE 2009 Chapter of the Year Award. Co-editor of two books by STAR, Springer Tracks in Advanced Robotics, Springer (2003, 2005). Research interests are in haptics, grasping, visual servoing, mobile robotics and geometric control. Author of more than 200 papers in those fields. Leader of a research unit in four EU projects: ROBOCAST, THE, ACTIVE, DALI. Coordinator of the EU ECHORD-EXPERIMENT HANDS.DVI, and of the IP collaborative project WEARHAP. He is a fellow of the IEEE.