# On equitably approaching and joining a group of interacting humans

Vishnu Karakkat Narayanan, Anne Spalanzani, François Pasteau, Marie Babel

# On equitably approaching and joining a group of interacting humans

Vishnu K.Narayanan[1], Anne Spalanzani[2], François Pasteau[3] and Marie Babel[3]

*Abstract*— In this work we introduce a low-level system that could be employed by a social robot like a robotic wheelchair or a humanoid, for approaching a group of interacting humans, in order to become a part of the interaction. Taking into account an *interaction space* that is created when at least two humans interact, a meeting point can be calculated where the robot should reach in order to equitably share space among the interacting group. We propose a sensor-based control task which uses the position and orientation of the humans with respect to the sensor as inputs, to reach the said meeting point while respecting spatial social constraints. Trials in simulation demonstrate the convergence of the control task and its capability as a low-level system for human-aware navigation.

## I. INTRODUCTION

Social robotics and human-robot interaction is increasingly becoming the forefront of current robotics research. Human-aware navigation forms an important part of social robotics where, for example, social constraints like avoiding the *personal* space of humans is explicitly considered [1] [2]. This paper aims to contribute specifically in the area of human-aware navigation, that is on meeting and joining groups in interaction. More precisely, this work demonstrates a control design for equitably joining groups in order to become a part of it while respecting spatial social constraints.

The core aim of human-aware navigation is to autonomously navigate and interact in a dynamic human populated environment, where social awareness is included in the navigation decisions [3]. Most of the literature available in this area focuses on how a robot can safely navigate *around* the personal spaces of humans, in order to avoid disturbance [3]–[5] or, in order to make room for human passing [6]. Moreover, complementary works [7]–[9] describe how a robot should initiate a conversation with a human being, where approach distances, gaze directions, the greeting process etc., are analysed in order to obtain solutions. But the key issue of *how to approach* humans with the intention of initiating conversation was initially tackled by Satake et. al. [9]. They designed a probabilistic path planning approach in order to frontally approach a single human target by taking into account the predicted trajectory of human target motion. As a further step, the present work deals with the problem of how to approach a group of humans in interaction, with the aim of becoming a part of the group.

In this context, preliminary studies conducted by Butler [10] indicated that an indirect pattern of approach by the
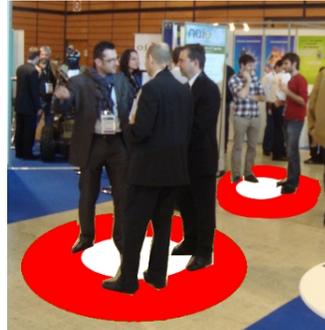


Fig. 1. People in conversation follow specific patterns of spatial arrangement. The situation of O-Space is marked with a white circle for the two groups and it is surrounded by the p-Space in red.

robot is typically considered as the favourite, as it decreases the threat of contact. In contrast recent studies have shown that a frontal approach is more desirable for humans and even more so in the case of groups [7], [9], [11]. But the problem of designing algorithms that can be utilized by a social robot in order to perform a frontal approach while respecting spatial social constraints is highly non-trivial. Works by Escobedo et. al. [12], [13] have defined an algorithm which is able to calculate meeting points for *detected* interacting groups. The said meeting point was reached by planning a trajectory using an $A^*$ algorithm and Dynamic Window path planner. In contrast, the present work illustrates a robust sensor-based (and sensor-agnostic) control system that is capable of reaching the optimal meeting point in a way that is socially conventional. We also argue that the system can be stacked onto a higher level controller for enhancing its capabilities.

The detection/perception part of this work is derived from the *Social Filter module* developed by Martinez [14]. The module facilitates the detection and representation human groups in interaction and can be used in conjunction with a laser scanner, an RGB-D sensor or even a stereo camera rig.

Thus the main contribution of this work is the design and verification of this sensor-based control system. The problem of approaching a specific group is broken down into two sub-tasks that have to be performed simultaneously. A stacking of these sub-tasks is performed based on a well-defined formalism [15] with some specific constraints that ensures the robot follows spatial social conventions. Analysis in simulation shows the efficiency the control system as a low-level module for human-aware navigation in social environments.

[1]Vishnu K.Narayanan is with the Lagadic Team at Inria Rennes, France. `vishnu.karakkat-narayanan at inria.fr`

[2] Anne Spalanzani is with UPMF Grenoble and Inria Grenoble, France

[3]François Pasteau and Marie Babel are with the Lagadic Team at Inria Rennes, INSA Rennes, France

## II. INTERACTION SPACES

A social robot such as a humanoid or a robotic wheelchair should respect the physical *space* that is created when two or more people join together to form a *focused* interaction. When a standing group of people agree to sustain a single focus of visual and cognitive attention, the resulting interaction is termed as a focused one [16]. Conversations are then said to be focused interactions which can be translated into a common shared space in the environment. In order to detect a focused interaction in an environment where several individuals are present, the manner in which the individuals position and orient themselves with respect to others is the key factor. The ideas of O-spaces and F-formations are at the center of this solution.

As shown in Figure 1, people interacting in groups follow some spatial arrangement patterns. This spatial arrangement is termed as an *Interaction Space*. The O-space in a focused interaction is the shared area reserved for the activity that is established by the specific group. Only participants have access to it, they protect it and others tend to respect it [17]. O-space varies depending on body size, posture, position and orientation of each participant during the activity. The p-space is the space surrounding the O-space which is used for the placement of the participant bodies and also personal belongings.

Now according to [18], the term F-formation is used to designate the system of spatial-orientation arrangement and postural behaviours that people create and maintain in order to sustain their O-space. The shape of the F-formation strongly depends on the number of people involved, the relationship among them, the group attentional focus and on relevant environmental constraints. For example when we consider two people in conversation, six formations are the most frequent: N-shape, Vis-a-vis, V-shape, L-shape, C-shape and side-by-side [17] [18]. Illustrations for four formations namely Vis-a-vis, L-Shape, C-Shape and V-Shape are given in Figure 2 for reference.

### A. Two people formations

A geometric representation of the interaction space (i.e the O-space and p-space) for two people formations can be extracted from the position and orientation of the humans (say with respect to a global map or with respect to a sensor frame). For two humans $H_1$ and $H_2$, the geometry of the O-spaces and the p-spaces with respect to the ground/floor plane is given in Figure 3 for various F-formations. $H_{12}$ is the center of the line that joins the two humans, $\phi_1$ and $\phi_2$ the orientation angles with respect to an arbitrary frame, $C$ is the center of the O-space and $V_i$ is the focus point of the interaction. The distance between the two humans is denoted by $D_H$. Judging the meeting point (i.e the point where a robot should place itself in order to become a part of this group) is fairly simple. We can see that ideally, the robot should position itself on a specific point, within the p-space where the line that joins $H_{12}$ and $V_i$ passes through. The robot should also be facing the focus point of the interaction which is $V_i$.



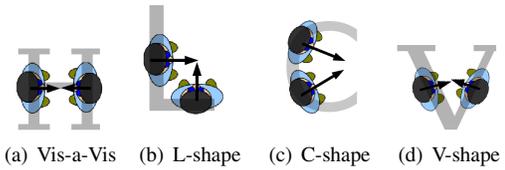(a) Vis-a-Vis  (b) L-shape  (c) C-shape  (d) V-shape

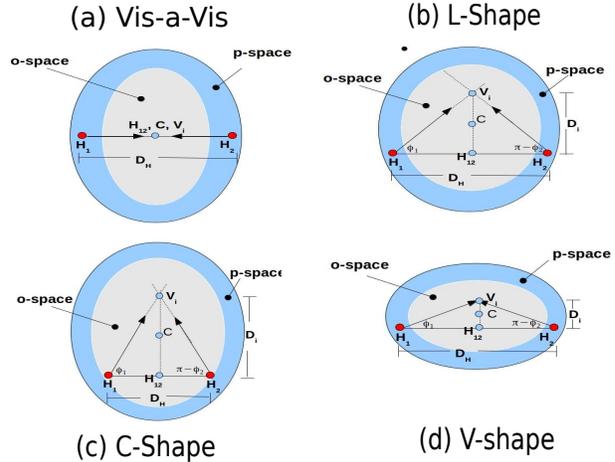Fig. 2. Frequent F-formations for 2 people groups



Fig. 3. The shape of the O-space and p-space with respect to the ground/floor plane for the 4 frequent F-formations in 2 people interactions. The geometrical relation between the participants is employed by the Social Filter module in order to represent Interaction Spaces.

### B. Formations with more than 2 people

In the case of formations with more than 2 people, the O-space is represented as a circle with the focus of attention located at the center of the circle. This phenomenon is more obvious as number of humans in the groups increases. With regards to calculating the meeting point, there can be a variety of solutions depending on the number, positions and orientations of the participating humans. For example, one solution would be to calculate the meeting point by considering the group as a 2 people formation with the 2 people who are farthest in terms of Euclidean distance.

A more detailed illustration on the detection of interaction spaces and F-formations as well as on meeting point calculation, as explained earlier, is detailed further in [3], [14]. It is not delineated in this study as it is not an objective of this work.

### C. Approaching a group

The following section (III) details the control design in order to equitably approach and join a detected group. Taking into account the fact that a frontal approach is the better solution [9], a robot should approach a group without entering the O-space of the interaction. Another point to consider is that a robot should reveal its intention of imminent approach to the group members. This can be achieved by ensuring the fact that the robot initially faces the group and at no point during the motion faces away from the group (i.e. the sensor always gazes at the group). With respect to the control design

presented in this work, for example, if an assistive wheelchair equipped with a laser scanner in the front is planning to join a group, it is advantageous to turn the wheelchair initially towards the group at the start of the motion and also ensure that the front of the wheelchair, at no point during the motion, faces away from the group. For a humanoid equipped with an RGB-D system or a stereo rig, it is advantageous to face the head towards the group as the humanoid moves to join the group.

## III. CONTROL SYSTEM

### A. Modelling

Considering a sensor frame $\mathcal{F}_s(P_S, x_s, y_s, z_s)$, we assume that the position and orientation of the detected humans and all geometric information regarding the interactions are available in this frame. We model the robot as a non-holonomic unicycle-type robot, which holds for systems such as robotic wheelchairs, while unicycle-type dynamics can be converted to walking motions for humanoids easily as demonstrated in [19]. Therefore we control two velocities namely the translational velocity $v$ and the angular velocity $\omega$. The robot frame is denoted by $\mathcal{F}_r(P_O, x_r, y_r, z_r)$. If the sensor is rigidly fixed so that we have a translation vector ${}^s\mathbf{t}_r = (w, 0, -l)$ between $\mathcal{F}_s$ and $\mathcal{F}_r$, we have the relationship between the velocity expressed in the robot frame $\mathbf{v} = [v, 0, 0, 0, 0, \omega]^T$ and the velocity expressed in sensor frame $\mathbf{v_s}$ as

$$\mathbf{v_s} = {}^s\mathbf{W}_r\mathbf{v}, \tag{1}$$

with ${}^s\mathbf{W}_r$ representing the velocity transformation matrix given by

$$ {}^s\mathbf{W}_r = \begin{bmatrix} {}^s\mathbf{R}_r & [{}^s\mathbf{t}_r]_\times {}^s\mathbf{R}_r \\ 0_{3\times3} & {}^s\mathbf{R}_r \end{bmatrix}. \tag{2}$$

${}^s\mathbf{R}_r$ in Eqn. (2) is the rotation matrix that models the fixed orientation of the sensor frame relatively to the robot frame and which is given by

$$ {}^s\mathbf{R}_r = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{bmatrix}. \tag{3}$$

Also, note that $[{}^s\mathbf{t}_r]_\times$ in Eqn. (2) represents the skew-symmetric matrix of ${}^s\mathbf{t}_r$.

### B. Task Features

In order to approach a specific interaction in an equitable way, we design a control task which facilitates a social robot reach an optimal meeting point by planning a socially acceptable spatial trajectory. Here we need to identify some features related to the group that can be exploited to design a control law that ensures that the robot reaches the said meeting point. We select the $x_s$ coordinate $X_h$ and the $z_s$ coordinate $Z_h$, with respect to the sensor frame, of the point representing the focus of attention, as the two initial task features. If the robot reaches the meeting point we observe that $Z_h$ should reach a desired value $Z_h^*$ that depends on the type of F-formation of the group (see Fig. 4). Whereas $X_h$
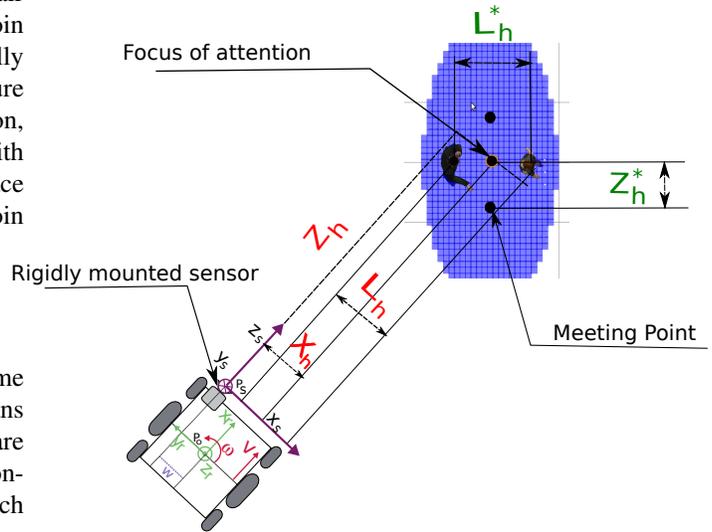


Fig. 4. A unicycle-type robot with respect to an interacting group. Robot and sensor frames along with the geometric representation of the task features are shown.

should attain a desired value of $X_h^* = 0$ which ensures that the sensor is aligned towards the group focus point $V_i$.

But we observe that regulating the errors $Z_h - Z_h^*$ and $X_h - X_h^*$ to zero does not ensure that the robot reaches the optimal meeting point. Therefore we introduce a new feature $L_h$ which is termed as the group length. In two-people formations, $L_h$ is the projection, on $x_s$, of the line segment connecting the two humans $H_1$ and $H_2$. In groups with more that two people $L_h$ is the projection, on $x_s$, of the line segment connecting the two farthest humans (in terms of Euclidean distance). The feature $L_h$ should reach a value of $L_h^* = D_H$ in two-people formations and $D_O$ (the diameter of the O-space) in groups with more that two people.

Therefore, in order to design the control law, we create a task $\mathbf{e}$ defined as

$$\mathbf{e} = [Z_h - Z_h^*, X_h - X_h^*, L_h - L_h^*]^T. \tag{4}$$

We can see that realizing the above task (i.e. regulating the errors $Z_h - Z_h^*$, $X_h - X_h^*$ and $L_h - L_h^*$ to zero) ensures that the robot reaches the said meeting point. In order to exponentially decrease the feature errors for the task $\mathbf{e}$, we can design a velocity controller as

$$\mathbf{v} = -\lambda \mathbf{J}_e^+\mathbf{e}, \tag{5}$$

where $\mathbf{J}_e^+$ represents the Moore-Penrose pseudo inverse of the task Jacobian $\mathbf{J}_e$ and $\lambda$ denotes the control gain.

Let $\mathbf{L} = [\mathbf{L}_{Z_h}, \mathbf{L}_{X_h}, \mathbf{L}_{L_h}]^T$ represent the interaction matrix that relates the dynamics of the task features $Z_h$, $X_h$ and $L_h$ with respect to the sensor velocity screw $\mathbf{v_s}$ such that $\mathbf{L}\mathbf{v_s} = \mathbf{e}$. The first two rows of the matrix $\mathbf{L}$ (i.e $\mathbf{L}_{Z_h}$ and $\mathbf{L}_{X_h}$, representing the features $Z_h$ and $X_h$) was determined specifically for the case of visual servoing by [20] and [21] for controlling a robot end-effector with respect to a 3-D image point while the third row ($\mathbf{L}_{L_h}$, representing

the feature $L_h$) was determined by [21] for visual servoing with respect to a line segment in the image. For the present case, the interaction matrix $\mathbf{L}$ takes up a form as follows

$$
\begin{aligned}
\mathbf{L} &= [\mathbf{L}_{Z_h}, \mathbf{L}_{X_h}, \mathbf{L}_{L_h}]^T \\
&= \begin{bmatrix} 0 & 0 & -1 & 0 & X_h & 0 \\ -1 & 0 & 0 & 0 & -Z_h & 0 \\ A & 0 & B & 0 & -L_h X_h & 0 \end{bmatrix}
\end{aligned}
\tag{6}
$$

,

where $A = \frac{Z_1 - Z_2}{Z_1 Z_2}$ and $B = \left( \frac{Z_1 + Z_2}{2 Z_1 Z_2} L_h - \frac{Z_1 - Z_2}{Z_1 Z_2} X_h \right)$.

Here $Z_1$ and $Z_2$ represent the $z_s$ coordinates of the positions of the two humans in a two-people formation. In a formation with more than three people, $Z_1$ and $Z_2$ represent the $z_s$ coordinates of the positions of the two farthest humans in terms of Euclidean distance.

As the robot Jacobian ${}^r\mathbf{J}_r$ expressed in the robot frame is equal to

$$
{}^r\mathbf{J}_r = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}^T,
\tag{7}
$$

we can finally define the task Jacobian $\mathbf{J}_e = \mathbf{L}\,{}^s\mathbf{W}_r\,{}^r\mathbf{J}_r$.

*C. Task Division and Stacking using Redundancy Formalism*

A simple look into the controller design given in Eqn. (5) would tell us that the the controller would be at most locally stable and would not perform well for the majority of cases as we regulate three features using two control variables. In order to ensure that the system is able to converge to a near global minima and, in addition, generate a motion that respects social constraints, we divide the above problem into 2 separate tasks: task 1 denoted by $\mathbf{e_1} = [Z_h - Z_h^*, X_h - X_h^*]^T$ and task 2 denoted by $\mathbf{e_2} = [L_h - L_h^*]$.

Task 1 can be intuitively seen as the robot moving forward in order to minimize the distance with respect to the group while simultaneously ensuring that the sensor gazes at the group. Whereas task 2 can be seen as a motion which ensures that the robot reaches the exact meeting point and is well aligned in order to face the group.

Here we observe that both the tasks are of full rank (i.e. they constrain both the DOFs of the robot). Thus it is essential to devise a method that facilitates the regulation of both the tasks simultaneously. Therefore we propose to activate (and/or partially activate) task 1 at specific intervals using an activation matrix. Consequently we can then project task 2 onto the null space of this new task 1 which means that task 2 can be activated (or partially activated) when task 1 is not activated (or partially activated). The activation matrix can be designed in such a way that the task priority is switched between task 1 and task 2 at optimum intervals so that the system converges to a near global minima. Moreover we can address the following issues while designing the said activation matrix.

1 The system should not fall into a local minima, particularly when the robot starts at a position as shown in Figure 5.
2 The robot should move in such a manner that respects the known social conventions described above.

3 The robot should not encroach the O-space if the meeting point is behind the humans in interaction.
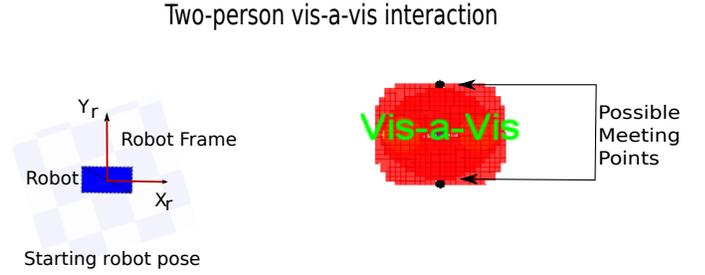
Two-person vis-a-vis interaction



Fig. 5. A robot starting position where there is a high possibility that the robot may encroach the O-space if it moves in order to join the group.

Thus, in order to modify the control law in Eqn. (5), the Jacobian $\mathbf{J}_e$ can be decomposed as

$$
\mathbf{J}_e = [\mathbf{J}_{Z_h}, \mathbf{J}_{X_h}, \mathbf{J}_{L_h}]^T = \begin{bmatrix} \mathbf{L}_{X_h}\,{}^s\mathbf{W}_r\,{}^r\mathbf{J}_r \\ \mathbf{L}_{Z_h}\,{}^s\mathbf{W}_r\,{}^r\mathbf{J}_r \\ \mathbf{L}_{L_h}\,{}^s\mathbf{W}_r\,{}^r\mathbf{J}_r \end{bmatrix}
\tag{8}
$$

.

Consequently the Jacobians for task 1 and 2 can be defined as

$$
\mathbf{J}_1 = \begin{bmatrix} \mathbf{J}_{Z_h} \\ \mathbf{J}_{X_h} \end{bmatrix}
\tag{9}
$$
$$
\mathbf{J}_2 = \mathbf{J}_{L_h}.
$$

If the activation function be denoted by $\mathbf{H}$, the velocity controller for the new task 1, which is activated at specific intervals, can be written as

$$
\mathbf{v} = -\lambda(\mathbf{H}\mathbf{J}_1)^+\mathbf{H}\mathbf{e_1},
\tag{10}
$$

where $\mathbf{H} = \mathrm{Diag}(h_{Z_h}, h_{X_h})$ is a diagonal matrix that weights the error $\mathbf{e_1}$ where $h_{Z_h}$ and $h_{X_h} \in [0;1]$ are the varying weights respectively associated to the features $Z_h$ and $X_h$. Such a concept of an interval that triggers the regulation of a task was illustrated in [23] and [24] within the domain of visual servoing for ultrasound tele-echography.

But it has to be noted that the matrix $(\mathbf{H}\mathbf{J}_1)^+\mathbf{H}$ does not remain continuous, as the rank of the matrix $\mathbf{H}\mathbf{J}_1$ may change from zero to 2 depending on the values of $h_{Z_h}$ and $h_{X_h}$. Therefore, in order to ensure a continuous control law, the pseudo-inverse operator $^+$ is replaced by the continuous pseudo-inverse operator $\oplus\,\mathbf{H}$ which was introduced in the framework of varying-feature-set [25]. The continuous inversion of the Jacobian $\mathbf{J}_1$ activated by the weight matrix $\mathbf{H}$ is given by

$$
\begin{aligned}
\mathbf{J}_1^{\oplus\mathbf{H}} =\ & h_{Z_h}(1 - h_{X_h}) \begin{bmatrix} \mathbf{J}_{Z_h} \\ \mathbf{0}_{1\times 2} \end{bmatrix}^+ \\
& + (1 - h_{Z_h})h_{X_h} \begin{bmatrix} \mathbf{0}_{1\times 2} \\ \mathbf{J}_{X_h} \end{bmatrix}^+ \\
& + h_{Z_h} h_{X_h} \mathbf{J}_1^+.
\end{aligned}
\tag{11}
$$

All the theoretical bases including the proof of continuity of this operator are detailed in [25].

Now, task 2 can be stacked on top of this varying rank task 1 so that task 2 is projected onto the null space of the new task 1 (i.e task 2 is activated when task 1 is not activated and vice versa). It can be achieved using the projection operator $\mathbf{P}_H = \mathbb{I}_2 - \mathbf{J}_1^{\oplus \mathbf{H}} \mathbf{J}_1$ as illustrated in [22]. According to [22], the controller will thus take up a form as follows:

$$\mathbf{v} = -\lambda[\mathbf{J}_1^{\oplus \mathbf{H}} \mathbf{e_1} - \mathbf{P}_H(\mathbf{J}_2 \mathbf{P}_H)^+(\mathbf{e_2} - \mathbf{J}_2 \mathbf{J}_1^{\oplus \mathbf{H}} \mathbf{e_1})]. \quad (12)$$

But, we also observe that the expression $\mathbf{P}_H(\mathbf{J}_2 \mathbf{P}_H)^+$ in Eqn. (12) is discontinuous over the whole task space. Therefore, this expression can be replaced by its continuous inversion $\mathbf{J}_2^{\mathbf{P}_H \oplus}$ where $\mathbf{J}_2^{\mathbf{P}_H \oplus} = ((\mathbf{J}_2^T)^{\oplus \mathbf{P}_H})^T$. The inversion $(\mathbf{J}_2^T)^{\oplus \mathbf{P}_H}$ is computed in the same fashion as described in Eqn. (11).

Consequently we can finally re-write the velocity controller (12) as

$$\mathbf{v} = -\lambda[\mathbf{J}_1^{\oplus \mathbf{H}} \mathbf{e_1} - \mathbf{J}_2^{\mathbf{P}_H \oplus}(\mathbf{e_2} - \mathbf{J}_2 \mathbf{J}_1^{\oplus \mathbf{H}} \mathbf{e_1})]. \quad (13)$$

Lastly, in order to eliminate the initially perceived failure cases we design the weight matrix $\mathbf{H} = \text{Diag}(h_{Z_h}, h_{X_h})$ and perform the task of joining a group using the following definitions:

$$h_{Z_h} = \begin{cases} 0.5 + \cos\left(\frac{\pi}{2}\frac{Z_h - \frac{2}{3}Z_{min}}{Z_{max} - \frac{2}{3}Z_{min}}\right) & \text{if } \frac{2}{3}Z_{min} \le Z_h \le Z_{max} \\ 1 & \text{if } \frac{1}{3}Z_{min} < Z_h < \frac{2}{3}Z_{min} \\ 0.5 - \cos\left(\frac{\pi}{2}\frac{Z_h - Z_h^*}{\frac{1}{3}Z_{min} - Z_h^*}\right) & \text{if } Z_h^* \le Z_h \le \frac{1}{3}Z_{min} \\ 0 & \text{otherwise} \end{cases} \quad (14)$$
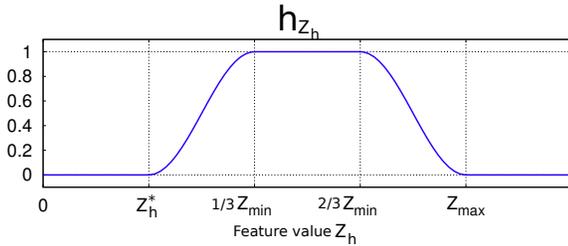
Fig. 6.   The evolution of $h_{Z_h}$ with respect to the feature $Z_h$

$$h_{X_h} = \begin{cases} 0.5 - \cos\left(\frac{\pi}{2}\frac{X_h - X_{min}}{X_{max} - X_{min}}\right) & \text{if } X_{min} \le X_h \le X_{max} \\ 0 & \text{if } -X_{min} < X_h < X_{min} \\ 0.5 - \cos\left(\frac{\pi}{2}\frac{X_h - X_{min}}{X_{max} - X_{min}}\right) & \text{if } -X_{max} \le X_h \le -X_{min} \\ 1 & \text{otherwise}. \end{cases} \quad (15)$$

Figure 6 visually represents the evolution of the activation factor $h_{Z_h}$ with respect to the parameters $Z_{min}$ and $Z_{max}$. $Z_{max}$ is defined as the maximum possible positive value $Z_h$ could obtain at a particular configuration. It can be obtained from the estimations of $X_h$ ($X_{init}$) and $Z_h$ ($Z_{init}$) at the first instant when $Z_h > 0$. Therefore $Z_{max} = \sqrt{X_{init}^2 + Z_{init}^2}$. Whereas $Z_{min} = Z_{max} - Z_h^*$. The above definition of $h_{Z_h}$ ensures that the activation of the feature $Z_h$ is 0 as long
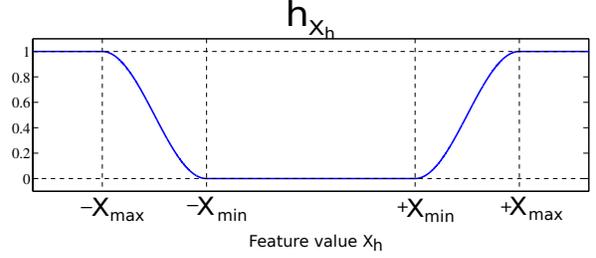
Fig. 7.   The evolution of $h_{X_h}$ with respect to the feature $X_h$

as $Z_h$ becomes positive. Therefore the robot performs a motion that regulates the errors on $X_h$ and $L_h$ which is essentially aligning the sensor towards the group (gazing at the group). It can serve as a notification of approach for the group members. As soon as $Z_h > 0$, the control over this feature is returned in a smooth fashion as the robot moves forward and towards the meeting point (see Figure 6 interval $[\frac{1}{3}Z_{min}, Z_{max}]$). Also, as the feature error $Z_h - Z_h^* \to 0$, the activation gradually decreases to zero which essentially allows the other two features to converge to their desired values (see Figure 6 interval $[Z_h^*, \frac{1}{3}Z_{min}]$).

Figure 7 represents the evolution of $h_{X_h}$ with respect to the intervals $[X_{min}, X_{max}]$ and $[-X_{min}, -X_{max}]$. This interval is designed as a dynamic interval which varies proportionally with the absolute value of the error of task 2 namely $|\mathbf{e_2}|$. Therefore $X_{min} = k_1|\mathbf{e_2}|$ and $X_{max} = k_2 X_{min}$. This variation ensures that the task priority can be switched from task 1 (when $|\mathbf{e_2}|$ is low) to task 2 (when $|\mathbf{e_2}|$ is high) at optimum intervals. If $|\mathbf{e_2}|$ is low, then the interval $[-X_{min}, X_{min}]$ is small which facilitates more regulation over the feature $X_h$ and when $|\mathbf{e_2}|$ is high, then the interval $[-X_{min}, X_{min}]$ is large which facilitates more regulation over the feature $L_h$. This also ensures that a local minima is avoided and the system reaches a near global minima. The value of the factors $k_1$ and $k_2$ can be tuned empirically so that the sensor is always gazing at the group during the motion.

The above definition of the matrix $\mathbf{H} = \text{Diag}(h_{Z_h}, h_{X_h})$ solves the first two of the three issue mentioned in Section IIIC. With respect to the third case, if the meeting point is behind the humans which means that the humans are facing away from the robot, the only way to solve this case is to move the robot in open loop till the meeting point is in front of the humans with respect to the robot (Nevertheless this is not an issue in an two person vis-a-vis formation).

It is also evident to note that the control law is capable of adapting to the dynamic motions of the humans within an interaction as long as the interaction does not break.

## IV. ANALYSIS

In order to verify the efficiency of the proposed control law as a suitable low-level system which can be integrated onto a social robot, we have devised scenarios in simulation to assess the trajectory of the robot. We simulate a non-holonomic type robot equipped with a laser scanner which
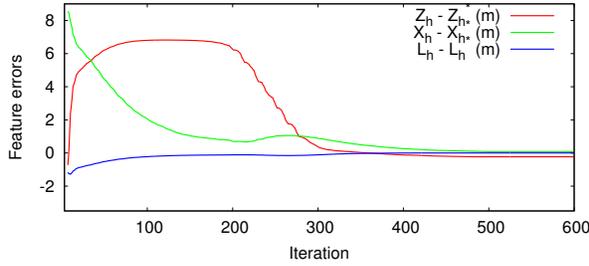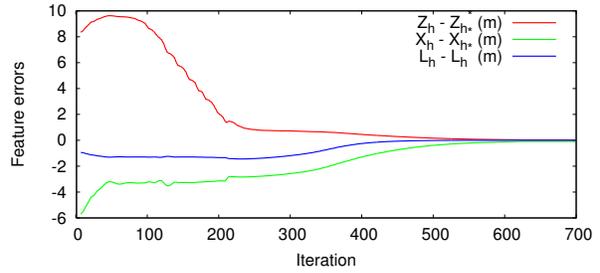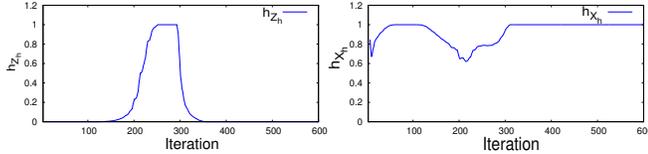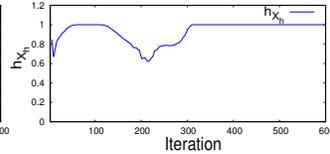
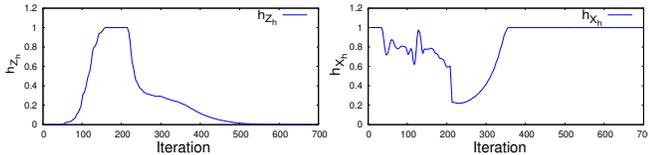Fig. 8. The feature errors $Z_h - Z - h^*$, $X_h - X_h^*$ and $L_h - L_h^*$ for Scenario 1 (a) and Scenario 2 (b)
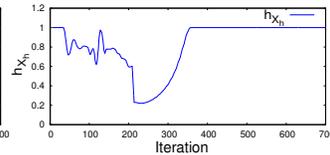


Fig. 9. The evolution of activation factors $h_{Z_h}$ and $h_{X_h}$ for Scenario 1



Fig. 10. The evolution of activation factors $h_{Z_h}$ and $h_{X_h}$ for Scenario 2



Fig. 11. Scenario 1: The initial position of the robot with respect to the humans $H_1$ and $H_2$ (a) and the trajectory taken by the robot in order to reach the meeting point (b)

can detect interactions using the Social Filter module. The simulations were carried out within the ROS platform while the estimation of the features and the control law computation was carried out using the ViSP [26] software. The gain $\lambda$ was set at $0.5$. The factors $k_1$ and $k_2$ can be tuned based on the field of view (FOV) of the sensor in such a way that the features does not leave the FOV. For the present simulations $k_1$ and $k_2$ were set at at $1.5$ and $2$ respectively.

We present two Scenarios for analysis. The robot initially is positioned as depicted in Figures 11(a) and 12(top) in Scenarios 1 and 2 respectively. In the first case, the robot is oriented away from a group of conversing humans ($H_1$ and $H_2$) in a V-shape interaction while it is positioned in front of the meeting point. The second case shows a starting configuration similar to the one presented in Fig. 5, where there is a high possibility of the control law falling into a local minima or encroaching the O-space. We simulate ideal cases where no obstacles (static or dynamic) are considered in the configuration in order to asses the performance of the control law.

The robot was tasked with joining the group using the proposed control scheme in both the Scenarios. The trajectories taken by the robot are presented in Figures 11(b) and 12(bottom). Figure 8 shows the feature errors ($Z_h - Z_h^*$, $X_h - X_h^*$ and $L_h - L_h^*$) between the task features and its
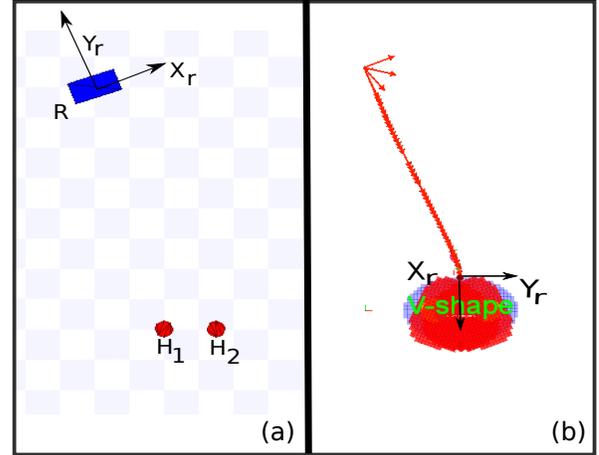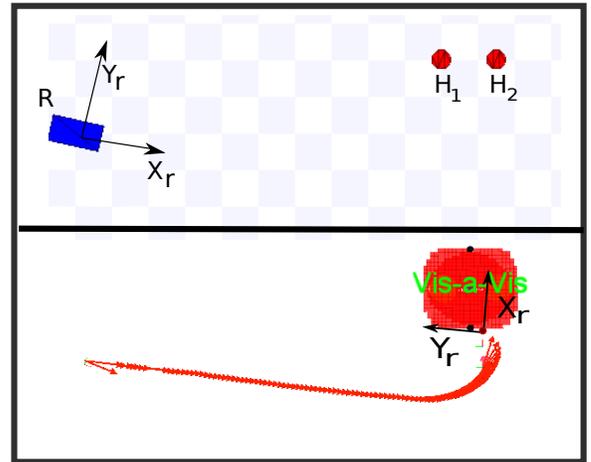


Fig. 12. Scenario 2: The initial position of the robot with respect to the humans $H_1$ and $H_2$ (top) and the trajectory taken by the robot in order to reach the meeting point (bottom)

desired values both the Scenarios while Figures 9 and 10 shows the evolution of the activation factors $h_{Z_h}$ and $h_{X_h}$ for Scenarios 1 & 2 respectively.

In the first Scenario the robot initially turns towards and faces the group, while regulating only $X_h$ and $L_h$. This can be observed from Figure 8(a) [Iteration 0 - 200] where $X_h$ rapidly converges to a low value (which is essentially the sensor gazing at the group). As the feature $Z_h$ is activated (see Figure 9a), it performs a natural trajectory towards the meeting point till it finally reaches the desired point.

Whereas in the second Scenario, we see that the robot moves in such a fashion such that it ensures the meeting point is reached without encroaching the O-space. Instead of moving straight, the motion is in a way that the robot moves in front of the group with a tolerance in order to turn towards the group meeting point. The evolution of the factor $h_{X_h}$ (Figure 10(b)) shows that priority is shared between task 1 (i.e regulation of $X_h$) and task 2 (i.e regulation of $L_h$) during the most of the motion in order to reach a near global minima. As the feature $Z_h \rightarrow Z_h^*$, the factor $h_{Z_h}$ drops off smoothly to zero which allows the other two features to converge to their desired values which can be observed from the final turn in the trajectory.

## V. Conclusion

We demonstrated a task-based control law which can serve as a low-level system for equitably joining interacting groups, while confirming to social conventions. The system uses the position and orientation of the participating humans with respect to a rigid sensor frame in order to control the translational and rotational velocity of a non-holonomic robot so that the robot positions itself aptly at the meeting point. A novel algorithm is designed which ensures that the motion adheres to social standards while ensuring that the controller does not fall into a local minima. Analysis in simulation illustrates the capability of the algorithm as a viable low-level system for human-aware interaction. Obstacle avoidance is not explicitly considered in this work since the control system can be easily fused with a higher level controller capable of performing the same. Future tests aim at evaluating the system on a social robot such as a robotic wheelchair in a real-world dynamic social environment.

## VI. Acknowledgment

## References

[1] D. Vasquez, P. Stein, J. Rios-Martinez, A. Escobedo, A. Spalanzani, and C. Laugier, "Human Aware Navigation for Assistive Robotics," *Proc. of International Symposium on Experimental Robotics* vol. 88, 2013.

[2] T. Kruse, A. K. Pandey, R. Alami, and A. Kirsch, "Human-aware robot navigation: A survey," *Robotics and Autonomous Systems*, vol. 61, no. 12, 2013, pp. 1726–1743.

[3] J. Rios-Martinez, A. Spalanzani, and C. Laugier, "Understanding human interaction for probabilistic autonomous navigation using Risk-RRT approach," in *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems*. 2011, pp. 2014–2019.

[4] E. Sisbot, A. Clodic, L. Marin U., M. Fontmarty, L. Brethes, and R. Alami, "Implementing a Human-Aware Robot System," in *Proc. of The 15th IEEE International Symposium on Robot and Human Interactive Communication*. 2006, pp. 727–732.

[5] E. Sisbot, R. Alami, T. Simeon, K. Dautenhahn, M. Walters, and S. Woods, "Navigation in the presence of humans," in *Proc. of 5th IEEE-RAS International Conference on Humanoid Robots*. 2005, pp. 181–188.

[6] E. Pacchierotti, H. Christensen, and P. Jensfelt, "Design of an Office-Guide Robot for Social Interaction Studies," in *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems*. 2006, pp. 4965–4970.

[7] K. Dautenhahn, M. Walters, S. Woods, K. L. Koay, C. L. Nehaniv, A. Sisbot, R. Alami, and T. Siméon, "How may I serve you?" in *Proc. of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*. 2006, p. 172.

[8] K. W. Strabala, M. K. Lee, A. D. Dragan, J. L. Forlizzi, S. Srinivasa, M. Cakmak, and V. Micelli, "Towards Seamless Human-Robot Handovers," *Journal of Human-Robot Interaction*, vol. 2, no. 1, pp. 112–132, Mar. 2013.

[9] S. Satake, T. Kanda, D. Glas, M. Imai, H. Ishiguro, and N. Hagita, "How to approach humans?-strategies for social robots to initiate interaction,"in *Proc. of the ACM SIGCHI/SIGART conference on Human-robot interaction*. pp. 109–116, 2009.

[10] J. T. Butler and A. Agah, "Psychological Effects of Behavior Patterns of a Mobile Personal Robot," *Autonomous Robots*, vol. 10, no. 2, pp. 185–202, Mar. 2001.

[11] D. Karreman, L. Utama, M. Joosse, M. Lohse, B. van Dijk, and V. Evers, "Robot etiquette," in *Proc. of the ACM/IEEE international conference on Human-robot interaction*. 2014, pp. 196–197.

[12] A. Escobedo, A. Spalanzani, and C. Laugier, "Multimodal control of a robotic wheelchair: Using contextual information for usability improvement," in *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems*. 2013, pp. 4262–4267.

[13] A. Escobedo, A. Spalanzani, and C. Laugier, "Using social cues to estimate possible destinations when driving a robotic wheelchair," in *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems*. 2014, pp. 3299–3304.

[14] J. Rios-Martinez, "Socially-Aware Robot Navigation: combining Risk Assessment and Social Conventions.", PhD Thesis, University of Grenoble, 2013.

[15] N. Mansard, O. Khatib, and A. Kheddar, "A unified approach to integrate unilateral constraints in the stack of tasks," *IEEE Transactions on Robotics* vol. 25, no. 3, pp. 670–685, 2009.

[16] Erving Goffman, *Behavior in Public Places: Notes on the Social Organization of Gatherings*. Free Press, 1966.

[17] A. Kendon, *Development of Multimodal Interfaces: Active Listening and Synchrony*, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, Mar. 2010, vol. 5967.

[18] T. M. Ciolek and A. Kendon, "Environment and the Spatial Arrangement of Conversational Encounters," *Sociological Inquiry*, vol. 50, no. 3-4, pp. 237–271, 1980.

[19] A. Faragasso, G. Oriolo, A. Paolillo, and M. Vendittelli, "Vision-based corridor navigation for humanoid robots," in *Proc. of IEEE International Conference on Robotics and Automation*. 2013, pp. 3190–3195.

[20] B. Espiau, F. Chaumette, and P. Rives, "A New Approach to Visual Servoing in Robotics," *IEEE Transactions on Robotics and Automation*, vol. 8, no. 3, 1992, pp. 313–326.

[21] F. Chaumette and S. Hutchinson, "Visual servo control, Part I: Basic approaches," *IEEE Robotics and Automation Magazine*, vol. 13, no. 4, 2006, pp. 82–90.

[22] N. Mansard and F. Chaumette, "Task Sequencing for High-Level Sensor-Based Control," *IEEE Transactions on Robotics*, vol. 23, no. 1, 2007,pp. 60–72.

[23] O. Kermorgant and F. Chaumette, "Combining IBVS and PBVS to ensure the visibility constraint," in *Proc. of IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*. 2011, pp. 2849–2854.

[24] T. Li, O. Kermorgant, and A. Krupa, "Maintaining visibility constraints during tele-echography with ultrasound visual servoing," in *Proc. of IEEE Int. Conf. on Robotics and Automation*.2012, pp. 4856–4861.

[25] N. Mansard, A. Remazeilles, and F. Chaumette, "Continuity of Varying-Feature-Set Control Laws," *IEEE Transactions on Automatic Control*, vol. 54, no. 11, 2009, pp. 2493–2505.

[26] E. Marchand, F. Spindler, and F. Chaumette, "ViSP for visual servoing: a generic software platform with a wide class of robot control skills," *IEEE Robotics and Automation Magazine*, vol. 12, no. 4, 2005, pp. 40–52.