<http://www.diva-portal.org>

Preprint

This is the submitted version of a paper presented at *The 28th IEEE International Conference on Robot and Human Interactive Communication – RO-MAN 2019, New Delhi, India, October 14-18, 2019..*

Citation for the original published paper:

Krishna, S., Kristoffersson, A., Kiselev, A., Loutfi, A. (2019)
Estimating Optimal Placement for a Robot in Social Group Interaction
In: *IEEE International Workshop on Robot and Human Communication (ROMAN)*
IEEE
<https://doi.org/10.1109/RO-MAN46459.2019.8956318>

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:oru:diva-78832>

Estimating Optimal Placement for a Robot in Social Group Interaction

Sai Krishna Pathi¹, Annica Kristofferson², Andrey Kiselev¹ and Amy Loutfi¹

Abstract—In this paper, we present a model to propose an optimal placement for a robot in a social group interaction. Our model estimates the O-space according to the F-formation theory. The method automatically calculates a suitable placement of the robot within a group of people. An evaluation of the method has been performed by conducting an experiment where participants stand in different formations and a robot is teleoperated to join the group. In one set of experiments, the operator positions the robot according to the specified location given by our algorithm. In another set of experiments, operators have the freedom to position the robot according to their personal choice. Follow-up questionnaires were performed to determine which of the placements were preferred by the participants. Our results indicate that the proposed method for automatic placement of the robot is supported from the view of the participants. The contribution of this work resides in a novel method to automatically estimate the best placement of the robot, as well as the results from user experiments to verify the quality of this method. These results suggest that teleoperated robots e.g. mobile robot telepresence systems could benefit from tools that assist operators in placing the robot in groups in a socially accepted manner.

I. INTRODUCTION

When a group of humans are interacting in the same physical space, they often form spatial formations in certain patterns, e.g. standing in a ring. These patterns described by the *F-formation* theory [6], [7] are said to arise “...whenver two or more people sustain a spatial and orientational relationship in which the space between them is one to which they have equal, direct and exclusive access.” (p. 209, [6]). Likewise, when humans join an existing group formation, they adhere to *F-formation* theory in order to promote successful interaction. As the use of social robots are increasing, it is equally important that such robots are able to relay on the same social behaviours to enable successful interaction with individuals as well as groups.

There are four F-formations and three social spaces within the F-formation theory. *O-space* is the convex empty space which all people involved in an interaction surround and are directed towards. *P-space* is the narrow strip on which people are standing while conversing and *R-space* is the space beyond the P-space. A *Vis-a-Vis* occurs when two people are interacting and facing each other. *L-shape* occurs when two people are standing perpendicular to each other, i.e., when they are situated on the two edges of the letter ‘L’. It is a *Side-by-Side* when the people interacting are standing close to each other and facing the same direction. When three

or more people are interacting within a circle, they are in a *Circular* formation.

In this paper, we present and evaluate a method to enable a robot to consider people’s spatial and orientational information in a group. This method then automatically proposes a placement for the robot that adheres to group formations. This location is referred to robot positioning spot (RPS). The method is based on the position and orientation of the people in the existing group (x, y, θ) and computes an optimal placement using the notion of the O-space (the convex empty space between a group of persons). Using the proposed approach, the method is evaluated with an experiment where a robot is teleoperated to a position in a group. In one condition, a human pilot joins the group according to his/her preference while in the other condition a human pilot joins the group based on the recommended position from RPS.

The paper is organized as follows: Section II gives an overview of previous works. Section III presents our algorithms for estimating the F-formations and a spot for the robot in the social interactions. Section IV presents the experiment performed to evaluate our algorithms. The results of the experimental validation are presented in Section V. Finally, the applicability of the method within HRI and conclude the paper with future works in Section VI.

II. RELATED WORKS

In the literature, several works have explored the ability to automatic detect groups using F-formation theory in vision-based systems. In [3], a Hough voting strategy is used to locate the O-space which in return provides the groups of conversing people. Considering F-formations a clustering problem, [5] built a graph model, in which each node of the graph is a person and the edges are the affinity between pair of people, to find the dominant set [5]. A similar model which considers the body, orientation and proxemics was proposed in [10]. Another work proposing a Hough voting approach for detecting groups is [12] which employs the weighted Boltzmann entropy for scoring group hypothesis. In [13], partly the same authors also proposed another approach known as Graph-Cuts for F-formation (GCFF). Experiments were conducted to detect groups in still images using proxemic information [13]. A game-theoretic framework was developed embedding the socio-psychological concept of F-formations and the biological constraints of social attention in [15]. The authors generated a frustum and computed affinity to extract the F-formations. A joint learning framework for the individuals’ heads, body orientations and F-formations in videos was proposed in [11]. Frustum of attention was used to extract features from

¹School of Science and Technology, Örebro University, 70281 Örebro, Sweden (sai.krishna, andrey.kiselev, amy.loutfi)@oru.se

²School of Innovation Design and Engineering, Mälardalen University, 722 20 Västerås, Sweden annica.kristoffersson@mdh.se

(a)

(b)

Fig. 1: (a) Sample image of people interacting in the scene. The scene contains the robot and the table with discussions cards and questionnaires on the table (b) The modified turtlebot with snacks

individuals and classify them accordingly such as associates, singletons and members of F-formations in [17]. A framework to find the social interactions between people in the scene through the Inter-Relation Pattern Matrix, computed by the three-dimensional visual field of a person in the scene, is proposed in [2]. Another work, [16] proposed detecting the F-formations based on lower body estimation obtained by tracking the position and orientation of people in a scene.

Only a handful of the current approaches have approached the problem from an egocentric viewpoint i.e. the point of view of the robot. In [4], a search-based method was proposed to recover the structure of the social interaction using multiple first person views. Both real world and simulation experiments were conducted. An influential work which proposed egocentric vision to detect groups through supervised clustering using people's head poses is [1]. The work resulted in a data set, EGO-GROUP, made publicly available with ground truth information.

The aforementioned methods consider the underlying idea behind F-formations and detect groups in a scene. However, they do not reflect upon the F-formation in which the groups are standing. One method estimating F-formations is [14]. The method detects and tracks humans as well as extracts social cues using a depth camera and a laser range finder. Social situations between people interacting are inferred in order for a robot to interact with the robot. In this work, the robot is provided with a task manager with predefined tasks such as "Approach to ask their needs or if they need any help". However, the work does not propose a spot for robot to join the social interaction.

While all these works have dealt with F-formations, none of them has provided a way or plan how to join groups. We assume that a spot or region is needed for a robot to join a group and provide assistance, engage in a social interaction and this is the focus of the work proposed here.

III. METHOD

Our method considers the spatial and orientation information of people in the group and estimates an RPS for the robot. The method calculates the intersecting point of people's gaze, i.e., the center of the O-space. Then, using the proxemics and F-formation theories, the method finds suitable empty spot/s on the P-space.

In this process, we assume that the following four parameters are known:

- 1) We know the distance between people and the robot.
- 2) We know the orientation of the people with respect to the robot.
- 3) Using both these information's transfer the egocentric view into a top-down view as in [1].
- 4) All people in the scene form the group (i.e. there is only one group in the scene and all people are in this group).

Due to fluctuating values of orientation of the person, it is assumed to be modelled to 4 distinct head poses possibly representing the four distinct facing directions of the person with respect to the robot. 0° when the person is not facing

the camera, i.e., when the back of his/her head is facing the camera. Then, proceeding in a clock-wise direction, the orientation is $\frac{\pi}{2}$ when the person is facing to the right. The orientation is π when the person is facing the camera and $\frac{3\pi}{2}$ when the person is facing to the left.

In this paper, we present an algorithm (1) which considers the number of people in the group and their relative orientation to classify them into F-formations.

Algorithm 1 Estimating F-formations

```

0: function PARTITIONGROUPSINTOF-FORMATIONS()
  Data: i is the first person, j is the second and k is the third person
   $\theta$  = orientation of the person
   $G$  = Number of people in the group
   $\Delta\theta$  : relative angle of  $\theta_i$  &  $\theta_j$  i.e.,  $(|\theta_i - \theta_j|)$ , modelled to  $0 \sim \pi$  i.e.,  $0, \frac{\pi}{2}, \pi$ 
1: if  $G = 2$  then
2:   if  $\Delta\theta = 0$  then
3:     Side-by-Side formation
4:   else if  $\Delta\theta = \frac{\pi}{2}$  : then
5:     L_shape formation
6:   else if  $\Delta\theta = \pi$  : then
7:     Vis-a-Vis formation
8:   else
9:     No Formation
10:  end if
11: else if  $G = 3$  then
12:   if  $(\Delta\theta_{ik} = \pi)$  OR  $(\Delta\theta_{jk} = \frac{\pi}{2})$  then
13:     Circular formation
14:   else
15:     No Formation
16:   end if
17: end if
17: end function=0

```

The algorithm (1) is based on the notion that the relative orientation of two people interacting would be similar to the way in which they are standing. When two people are standing in a Vis-a-Vis formation, it appears as if they are standing on a straight line. Similarly, when two people are standing in an L-shape, it appears as if they are standing in a right-angled triangle. In Side-by-Side formations, people are facing the same direction. i.e., their orientation is the same and $\Delta\theta = \text{abs}|\theta_i - \theta_j|$ is 0° .

The estimated F-formation is then used to find the RPS for joining the group. The aim is to detect the O-space between people in the group and then find an empty location on the group's P-space by considering the spatial and orientation information of the people in the group. The RPS varies between the F-formations and a global model cannot be proposed for all the formations. We calculate the RPS using a specific algorithm depending on the estimated F-formation.

If two people are interacting in a Vis-a-Vis, L-shape or Side-by-Side formation, the person standing on the left with respect to the robot is denoted by i while the person to the

right is denoted by j . The algorithms for the three formations are presented below.

For **Vis-a-Vis formation**, we calculate the distance between the people. The radius r of the O-space is $d/2$. Using r and the relative angle (θ) of the people, we obtain the coordinates of the center (origin) of O-space. Imagine a line $L1$ between i and j , draw a line $L2$ of length d perpendicular to $L1$ in such a way that the mid points of both the lines intersect. The endpoints of the line $L2$ proposes the RPS for joining the interaction. These RPS are in the P-space and on the edge of the O-space and the robot should be oriented towards the origin (see algorithm(2) and Fig. 2a).

Algorithm 2 Vis-a-Vis formation

```

0: function RPS_VIS-A-VIS_F-FORMATION()
  Data: i and j are in interaction
   $\theta_i$  = orientation of i
1: if Vis-a-Vis formation then
2:    $d = \text{Euclidean}(i, j)$ 
3:   radius of O-space ( $r$ ) =  $d/2$ 
4:   if  $\theta_i = \frac{\pi}{2}$  : then
5:      $i(x+r, y) \rightarrow$  origin of O-space,  $O(x, y)$ 
6:      $RPS = O(x, y+r)$  and  $O(x, y-r)$ 
7:   else if  $\theta_i = 0$  : then
8:      $i(x, y+r) \rightarrow$  origin of O-space,  $O(x, y)$ 
9:      $RPS = O(x+r, y)$  and  $O(x-r, y)$ 
10:  else if  $\theta_i = \pi$  : then
11:     $i(x, y-r) \rightarrow$  origin of O-space,  $O(x, y)$ 
12:     $RPS = O(x+r, y)$  and  $O(x-r, y)$ 
13:  end if
14:   $RPS_\theta$  = facing towards the O-space origin
15: end if
15: end function=0

```

For **L-shape Formation**, the intersecting point of the people's gaze is calculated by drawing a line in the direction of orientation of each person. Solving the mathematical equations for line segments provides us with the origin of the O-space. More simplified, one line is perpendicular to the X-axis and other line is parallel to the X-axis. The origin of the O-space, i.e., the intersecting point, $Ip(x, y)$ is x from the perpendicular line's starting point and y from the starting point of the parallel line. Then, we calculate the radius r between Ip and a person in the group. We obtain the RPS by adding or subtracting r to the origin coordinates on the X- and Y-axis respectively. The main idea of algorithm (3) is that Ip of the gaze directions of two people (i and j) involved in an interaction is the origin of the O-space. We can map i, j and Ip as a right-angled triangle and project and reverse it from the origin of the O-space. As shown in Fig. 2b, the edges of the reversed right-angled triangle are the RPS.

For **Side-by-Side Formation**, we calculate the distance d between two people instead of the distance to the origin of the O-space. Then, we obtain the RPS by adding or subtracting the distance to the people's coordinates based on the orientation of the group members (see algorithm (4) and Fig. 2c).

Algorithm 3 L-shape formation

```
0: function RPS_L-SHAPE_F-FORMATION()
  Data: i and j are in interaction
   $\theta_i$  = orientation of i
  L1 = Line between i and j
  L2 = Line perpendicular to L1 and midpoints of L1, L2 intersect
1: if L-shape formation then
2:   Draw a line for each person in their gaze direction
   and this lines will intersect each other
3:   Intersecting point ( $I_p$ ) = solve the line equations
4:    $y = ax + c$ ,  $y = bx + d$ 
5:    $x = d - c/a - b$ ,  $y = a(d - c/a - b) + c \rightarrow I_p(x, y)$ 
6:   radius of O-space ( $r$ ) = Euclidean( $I_p$ , i)
7:   if  $\theta_i = \frac{\pi}{2}$  : then
8:      $RPS = I_p(x + r, y)$ 
9:   else if  $\theta_i = 0$  or  $\pi$  : then
10:     $RPS = I_p(x - r, y)$ 
11:   end if
12: end if
12: end function=0
```

Algorithm 4 Side-by-Side formation

```
0: function RPS_SIDE-BY-SIDE_F-FORMATION()
  Data: i and j are in interaction
   $\theta_i$  = orientation of i
1: if Side-by-Side formation then
2:    $d = \text{Euclidean}(i, j)$ 
3:   if ( $\theta_i = 90$  or  $270$ ) and ( $i(y) > j(y)$ ) : then
4:      $RPS = i(x, y + d)$ 
5:   else if ( $\theta_i = 90$  or  $270$ ) and ( $i(y) < j(y)$ ) : then
6:      $RPS = i(x, y - d)$ 
7:   else if  $\theta_i = 0$  or  $180$  : then
8:      $RPS = i(x - d, y)$ 
9:   end if
10: end if
10: end function=0
```

For **Circular Formation**: If three people, i, j and k are standing in a circular formation. This pattern between adjacent individuals can be considered a hybrid of the Side-by-Side and the L-shaped formations [14]. In our case, the limited number of people in this formation is three, as the people increase and there will be no free spot left in the group. The L-shaped formation between two individuals is detected and then RPS is estimated which can be seen in figure 2d and the algorithm (5).

IV. EVALUATION

In the experiment, a group of two people (one participant and one confederate), were asked to stand in the scene and discuss three different topics. The confederate knows about F-formations but not about our approach and RPS. The confederate influenced the formation such that the group interacted in one formation (Vis-a-Vis, L-shape and Side-by-

Fig. 2: RPS for joining the F-formations.

Algorithm 5 Circular formation

```
0: function RPS_CIRCULAR_F-FORMATION()
  Data: i, j and k are in interaction
   $\theta_i, \theta_j, \theta_k$  = orientation of i, j, k
1: if Circular formation then
2:   if  $\theta_i = \frac{\pi}{2}$  and ( $\theta_j = 0$  or  $\pi$ ) : then
3:     i & j are in L-shape formations
4:     Use L-shape formation to find  $I_p$ , r
5:     if  $\theta_j = 0$  and  $\theta_k = \frac{3\pi}{2}$  : then
6:        $RPS = I_p(x, y + r)$ 
7:     else if  $\theta_j = \pi$  and  $\theta_k = \frac{3\pi}{2}$  : then
8:        $RPS = I_p(x, y - r)$ 
9:     else if  $\theta_k = 0$  or  $\pi$  : then
10:       $RPS = I_p(x + r, y)$ 
11:     end if
12:   else if ( $\theta_i = 0$  or  $\pi$ ) and  $\theta_k = \frac{3\pi}{2}$  : then
13:     i & k are in L-shape formation
14:     Use L-shape formation to find  $I_p$ , r
15:     if ( $\theta_j = 0$  or  $\pi$ ) : then
16:        $RPS = I_p(x - r, y)$ 
17:     end if
18:   end if
19: end if
19: end function=0
```

Side) per discussion topic. At two occasions per discussion, a robot joined the group. There were two different pilots, who were trained to teleoperate the robot in a professional way, joining the group during each discussion.

- Pilot_1 (P1) had no prior information about our method.
- Pilot_2 (P2) had knowledge about the proposed method for determining the RPS.

No predetermined order was used in the experimentation. In other words, the order of the group formations were randomized as well as the order of the pilots. Each experimental session took approximately 25 minutes.

The independent variables (conditions) in our experiment were:

- 1) The formation prior to the robot joining the group (L-shape, Vis-a-Vis, and Side-by-Side)

The ranking's provided by both the confederate and participants are gathered. These are based on the participants assessment of the appropriateness of the position of the robot in the groups. The appropriateness of the position is ranked on a scale from 1 to 7 where 7 is judged as highly appropriate and 1 is judged as highly inappropriate.

The dependent variables are:

- 1) The participants ranking of the placement of the robot.
- 2) The confederate's ranking of the placement of the robot as he/she would be ranking the robot's position based on the F-formation's theory.

We assume that the RPS approach would be considered more appropriate than teleoperator's placement of robot in the group by the group members. Therefore, the following hypothesis was postulated for this study and the dependent variable was selected accordingly.

- H1 **Participants would rank Pilot_2's placement higher than, or equal to, Pilot_1's in all the formations**
- H2 **Confederate would rank Pilot_2's placement higher than, or equal to, Pilot_1's in all the formations**

A. The Scene

The experiment took place in a robot lab which is equipped with a "Qualisys" system. The system provides cameras and software for capturing motion by precision and a 3D position tracking system. The system includes ten synchronized wall-mounted cameras and helmets with reflectors which are calibrated to provide accurate positional and orientation information in the lab area.

On the table, there were three down-facing cards, two piles with three forms and two pens. On each card, a current affair's discussion topic was presented. Each questionnaire contained two items. *Rank the first/second pilot's placement of the robot.* Both items were to be answered on a Likert scale 1-7 where 1 = very inappropriate to 7 = very appropriate. Nearby the table, there was an empty space in which the discussions take place as shown in the Figure 1a.

The robot used in the evaluation was a modified version of the Turtlebot. The robot is equipped with sensors allowing

the Qualisys system to acquire its positional and orientational information. Different types of snacks were placed in a box on the top of the robot with a note, "Do Help yourself". Before the start of the interaction, and in between all robot encounters, the robot was standing still in a pre-defined space as shown in Figure 1b.

Towards the side of the lab, there were two wall-facing working stations for the pilots. P1 had access to the robot's camera, which provided an egocentric view of the scene, which was used to join the groups. P2 joined the group in the RPS provided by our method using Qualisys software.

B. Experimental Procedure

The experiment started with welcoming the participants and thanking them for attending the experiment. Then, the participants were provided with the following description of the experiment. *The task is that both of you put on the helmets and there are three cards with a topic on each of them. Pick up a single card, go into the scene and start discussing the topic for approximately 5 minutes. While discussing, the robot will approach and join the group. There are a few snacks in the box on the robot, have them if you intend to. After 10-15 seconds, the robot will leave the group. After sometime, the robot will again join the group for the second time. Again, pick up and have the snacks if you want to. After 10-15 seconds the robot will leave the group. Then, you could conclude the discussion and come to the table and pickup the questionnaire. Rate the position of the robot in the group for the two times, for the first and second time. How appropriately the robot was positioned in the group with respect to both of you. Continue this process until you have discussed three topics and filled out three questionnaires. The helmet, and a number of wall-mounted cameras will be used for conducting the experiment. The cameras do not record any images but only tracks the helmets in the scene. The experiment, however, is anonymous and we will debrief you on how the information is used after concluding the experiment. Now, please provide your consent by signing this form and fill out the socio-demographic questionnaire.* The socio-demographics questionnaire gathered information on age and gender.

The below procedure was used for each group entering the room. Here we use the terms *first pilot* and *second pilot* for scripting purposes. These are not to be confused with the terms P1 and P2, i.e., the order in which the pilots joined the groups varied.

- 1) The group enters the room and walks to the table.
- 2) The group members fetches the top card.
- 3) The group moves into the scene and discusses the topic in the confederate's choice of F-formation.
- 4) After 30s, the first pilot teleoperates the robot and joins the group.
- 5) If interested, the participant takes some snacks.
- 6) The first pilot withdraws from the group with the robot after 20s.
- 7) Sometime into the discussion, the second pilot teleoperates the robot and joins the group.

- 8) Again, if interested, the participant takes some snacks.
- 9) The second pilot withdraws the robot from the group after 20s.
- 10) Then group members conclude the discussion and go back to the table.
- 11) The group members fill out the top questionnaire.
- 12) Step 2-11 is repeated twice after which the group is told to exit the room and a researcher provides them with a debriefing note upon exit.

The debriefing note contained the following information: *The goal of the experiment was to compare our algorithm's RPS with that of a teleoperator's choice of spot for joining the group. Thanks again for participating in this study.*

After each participant left the room, the participant's and confederate's questionnaires were collected and the Qualisys-recording was saved and exported into a csv-file.

V. RESULTS

In order to systematically evaluate our proposed method for determining the RPS, we conducted a within-subject experiment. The results are presented in this section.

A. Participants

There were 21 participants in the study (14 (66.7%) were male, 6 (28.57%) were female, and one person who preferred not to inform on the gender. The average age was $\mu=30.15$, $\sigma=5.23$ (age range: 21-41). All the participants were exposed to all experimental conditions.

B. Pilot's placement of robot

Using the RPS determined by the algorithm, Pilot_2 (P2) would position the robot on the P-space. Pilot_1 (P1) joined the groups without information about the existence of F-formations, O-space, P-space and R-space.

Using the Qualisys data, the pilots' positioning of the robot with respect to the group was extracted. Few instances of the robot positioned by P1 and P2 can be seen from Figure 3. While P1 was adjusting the position of the robot to join the group, P2 placed the robot on the edge of O-space i.e., P-space as seen in figure 3b. The robot has entered the O-space of the group while interacting in the L-shape which can be seen in figure 3c whereas the counter part P2 placed the robot on the P-space as seen in figure 3d. In Side-by-Side formation as seen in figure 3e, while joining the group, P1 hit a group member's foot and the person moved back and the robot stood in his/her place. All the time, in this formation, P1 was placing the robot in similar fashion i.e., in-front of a group member to join the interaction. This has a risk of robot exploiting the group member's personal space and also hitting the person which was the case. P2 placed the robot beside one of the group member as seen in figure 3f.

Every participant performed one experiment i.e., was exposed to three different formations. In two different experiments, three formations are not considered. Once in Side-by-Side formation, a wire was stuck in robot's wheel resulting in P2 being unable to move the robot forward (this position was ranked with a 1). P2 left the robot in that position to

Fig. 3: P1 and P2 joining the different formations. The blue and pink circle balls with green and red arrows represent members of the groups, having an interaction in different formations. The orange circle balls with red and green arrows represent the robot. The blue circle in-front of group members represent the O-space. (a) The overlap of the pink lines with orange lines indicate a collision between the robot and group member. (e) The overlap of lines indicate the collision.

rank which was not RPS. In order not to disturb the ongoing experiment. Twice in L-shape formation, first, the robot's battery died while conducting the last trial during the day. Second, there was a sudden failure in the tracking system for one of the helmets leading to a wrongly calculated RPS.

C. Participant's subjective ranking of robot's placement

To investigate H1, the response to the question *Rank the pilot's placement of the robot in the group* for each of the three formations and two pilots was analyzed. As shown in Table I, the participants rated both pilots' placement of the robot as appropriate although there were a small number of participants providing low ranks (see Fig. 4).

Four One-way ANOVA tests for the two pilots' placement of the robot were conducted (Vis-a-Vis, L-shape, Side-by-Side and all samples) based on participant's ranking. None of the tests showed a significant difference ($p < 0.05$) between

Fig. 4: Distribution of the participants and confederates' subjective ranking for the two pilots' placement of the robot when joining pairs standing in the three F-formations on a 1-7 likert scale where 1 = very inappropriate and 7 = very appropriate.

the pilots. Vis-a-Vis: $F(1,40) = 1.60$, $p < 0.21$, L-shape: $F(1,36) = 0.46$, $p < 0.50$, Side-by-Side: $F(1,38) = 1.41$, $p < 0.24$, and All: $F(1,118) = 3.32$, $p < 0.07$.

Hence, the experimental results performed using the participant's ranking does not show any significance about P2's placement of the robot using our approach compared to P1's placement of the robot.

TABLE I: Mean and standard deviation for the participant's subjective ranking of the two pilots' placement of the robot. n indicates number of participants.

F-formation	Pilot	μ	σ	n
Vis-a-Vis	P1	4.71	1.34	21
Vis-a-Vis	P2	5.23	1.33	21
L-shape	P1	5.31	1.63	19
L-shape	P2	5.63	1.21	19
Side-by-Side	P1	4.80	1.77	20
Side-by-Side	P2	5.45	1.70	20

D. Confederate's ranking of robot's placement

To investigate H2, the response to the question *Rank the pilot's placement of the robot in the group* for each of the three formations and two pilots was analyzed from the confederate's perspective. The confederate has knowledge about the F-formation's theory and applies this information while ranking the robot's position as how appropriately the robot was standing with respect to the group i.e., was the robot standing on the P-space or R-space or in O-space. From the Table II and figure 4, it is shown that our approach was rated higher (5,6 or 7) more number of times compared to P1's placement of the robot into the group except for Side-by-Side formation.

Four One-way ANOVA tests for the two pilots' placement of the robot were conducted (Vis-a-Vis, L-shape, Side-by-Side and all samples) based on confederate's ranking. Two of the tests showed a significant difference ($p < 0.05$) between the pilots which are L-shape and all samples. Vis-a-Vis: $F(1,40) = 2.59$, $p < 0.11$, L-shape: $F(1,36) = 5.87$, $p < 0.02$, Side-by-Side: $F(1,38) = 0.02$, $p < 0.88$, and All: $F(1,118) = 5.47$, $p < 0.02$.

Hence, the experimental results performed using the ranking collected from the confederate shows a mixed results. The results show significance in L-shape formation i.e., P2's position of the robot in the RPS was more appropriate than P1's position in L-shape formation.

TABLE II: Mean and standard deviation for the Confederate's ranking of the two pilots' placement of the robot. n indicates number of participants.

F-formation	Pilot	μ	σ	n
Vis-a-Vis	P1	6.28	1.58	21
Vis-a-Vis	P2	6.85	0.35	21
L-shape	P1	5.68	1.70	19
L-shape	P2	6.68	0.58	19
Side-by-Side	P1	6.5	1.10	20
Side-by-Side	P2	6.45	0.99	20

Finally, the present results must be interpreted cautiously and considered as preliminary results in this direction of estimating a spot for robot to join the social group interaction. Previously, researchers proposed to approach the group but not to place the robot on a specific spot in the group. This is first of its kind proposing a Robot Positioning Spot (RPS). The participant's ranking does not show any significance but confederate's ranking does show significance in L-shape formation. Disassembling and considering the results formation by formation would give an in-depth view into the results. In here, the ranking of 1,2,3 or 4 can be viewed as lower ranking and 5,6 or 7 can be viewed as higher ranking. Firstly Vis-a-Vis formation, 10 participants ranked P1's placement of the robot as lower and 11 participants as higher. P2's placement was ranked as lower by 6 participants and higher by 15 participants. More number of participants ranked P2's placement as higher compared to P1's placement. Confederate ranked P1's placement as lower in 2 experiments and higher in 19 experiments. P2's

placement as higher in all 21 experiments. Confederate also ranked P2's placement higher more number times than to P1. This indicates, according to participant and confederate, P2's placement was more appropriate in the group more number of times compare to P1's placement. Secondly, L-shape formation, participants ranked similarly for both the pilot's placement. This indicates participants found both the pilot's placement of robot approximately equal. Confederate narrates a totally different story regarding both the pilot's placement of the robot in the group. P1's placement was ranked as lower in 4 experiments and higher in 15 experiments. P2's placement was ranked as higher in all 19 experiments. This clearly shows that confederate considered P2's placement to be more appropriate more number of times than P1's placement. In both these formations, i.e., Vis-a-Vis and L-shape, in few occasions the robot entered the O-space to join the interaction. This is when the confederate ranked 1, 3, 4 and 5 based on how far the robot was into the O-space.

VI. OBSERVATIONS AND CONCLUSIONS

In HRI, the robot should be flexible and adapt to the social group interactions and position itself in the group to be part of the interaction. The social group interactions are structured based on different factors such as culture, gender, status, age, familiarity, relationship, pose, etc. [8], [9]. The O-space of the group varies due to this factors i.e., the O-space is smaller or larger. While joining the group, the robot should adopt the O-space and stand closer or farther to the members of the group. This would make the robot part of the group and not as unconventional for the members of group. During the study in many experiments the group members stood closer and little farther in others. Our approach was able to adapt the O-space of the group and propose the RPS accordingly. P2 joined the group but the same could not be said for the P1. This can be one of the reason why P1 entered the O-space in L-shape formation which resulted in significant result in L-shape through confederate's ranking.

Another important observation during the experiment was P1 took more time to join the groups. This was due to two main reasons, one is latency problem and the second being the placement of the robot in the group.

In this paper, we developed a model to propose a robot positioning spot (RPS) in a social group interacting in the scene. Our model estimates the O-space using the social cues proxemics and F-formation theories. The method calculates a suitable empty spot/s on the P-space. To evaluate our model systematically, we conducted an experiment which includes 21 participants standing in different formations and two pilots teleoperating a robot to join the social group interaction. One pilot joins the group in his/her choice of spot and other pilot uses our approach and places the robot in RPS. The group members (one is participant and other is confederate) rank the positions of the robot in the group. How appropriately the robot was positioned in the group with respect to the group members. Analyzing the experimental results demonstrate the effectiveness of the

proposed method and further this could be extended to make the robot automatically join the social group interactions.

ACKNOWLEDGMENT

Örebro University is funding the research through Successful Ageing Programme. The statements made herein are solely the responsibility of the authors. We would like to acknowledge our colleague André Potenza who did pilot the robot as Pilot_1(P1).

REFERENCES

- [1] Alletto, S., Serra, G., Calderara, S., Cucchiara, R.: Understanding social relationships in egocentric vision. *Pattern Recognition* **48**(12), 4082–4096 (2015)
- [2] Bazzani, L., Cristani, M., Tosato, D., Farenzena, M., Paggetti, G., Menegaz, G., Murino, V.: Social interactions by visual focus of attention in a three-dimensional environment. *Expert Systems* **30**(2), 115–127 (2013)
- [3] Cristani, M., Bazzani, L., Paggetti, G., Fossati, A., Tosato, D., Del Bue, A., Menegaz, G., Murino, V.: Social interaction discovery by statistical analysis of f-formations. In: *BMVC*, vol. 2, p. 4 (2011)
- [4] Gan, T., Wong, Y., Zhang, D., Kankanhalli, M.S.: Temporal encoded f-formation system for social interaction detection. In: *Proceedings of the 21st ACM international conference on Multimedia*, pp. 937–946. ACM (2013)
- [5] Hung, H., Kröse, B.: Detecting f-formations as dominant sets. In: *Proceedings of the 13th international conference on multimodal interfaces*, pp. 231–238. ACM (2011)
- [6] Kendon, A.: 7 Spatial organization in social encounters: the F-formation system. In: *Conducting interaction: patterns of behavior in focused encounters*, pp. 209–238. Cambridge University Press, Cambridge (1990)
- [7] Kendon, A.: Spacing and orientation in co-present interaction. In: V.C.H.A.N.A. Esposito A. Campbell N. (ed.) In *Proceedings of Development of Multimodal Interfaces: Active Listening and Synchrony*, Second COST 2102 International Training School. Lecture Notes in Computer Science, vol. 5967 LNCS, pp. 1–15. Springer, Berlin, Heidelberg (2010). DOI 10.1007/978-3-642-12397-9{_}1
- [8] Michalowski, M.P., Sabanovic, S., Simmons, R.: A spatial model of engagement for a social robot. In: *9th IEEE International Workshop on Advanced Motion Control*, 2006., pp. 762–767. IEEE (2006)
- [9] Patterson, M.: Spatial factors in social interactions. *Human Relations* **21**(4), 351–361 (1968)
- [10] Pavan, M., Pelillo, M.: Dominant sets and pairwise clustering. *IEEE transactions on pattern analysis and machine intelligence* **29**(1), 167–172 (2007)
- [11] Ricci, E., Varadarajan, J., Subramanian, R., Rota Bulò, S., Ahuja, N., Lanz, O.: Uncovering interactions and interactors: Joint estimation of head, body orientation and f-formations from surveillance videos. In: *Proceedings of the IEEE International Conference on Computer Vision*, pp. 4660–4668 (2015)
- [12] Setti, F., Lanz, O., Ferrario, R., Murino, V., Cristani, M.: Multi-scale f-formation discovery for group detection. In: *2013 IEEE International Conference on Image Processing*, pp. 3547–3551. IEEE (2013)
- [13] Setti, F., Russell, C., Bassetti, C., Cristani, M.: F-formation detection: Individuating free-standing conversational groups in images. *PloS one* **10**(5), e0123783 (2015)
- [14] Tseng, S.H., Chao, Y., Lin, C., Fu, L.C.: Service robots: System design for tracking people through data fusion and initiating interaction with the human group by inferring social situations. *Robotics and Autonomous Systems* **83**, 188–202 (2016)
- [15] Vascon, S., Mequanint, E.Z., Cristani, M., Hung, H., Pelillo, M., Murino, V.: A game-theoretic probabilistic approach for detecting conversational groups. In: *Asian conference on computer vision*, pp. 658–675. Springer (2014)
- [16] Vázquez, M., Steinfeld, A., Hudson, S.E.: Parallel detection of conversational groups of free-standing people and tracking of their lower-body orientation. In: *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 3010–3017. IEEE (2015)
- [17] Zhang, L., Hung, H.: Beyond f-formations: Determining social involvement in free standing conversing groups from static images. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1086–1095 (2016)