

# Join the Group Formations using Social Cues in Social Robots

Doctoral Consortium

Sai Krishna Pathi

Center of Applied Autonomous Sensor Systems (AASS)

Örebro University

Örebro, Sweden

sai.krishna@oru.se

## ABSTRACT

This work investigates how agents can spatially orient themselves into formations which provide good conditions for enabling social interaction. To achieve this, we are using socio-psychological notion, F-formation in our project and based on this concept, we detect positions of other agents in a scene to find the optimum placement. Using both simulation and real robotic systems, the system aims to achieve a functionality which enables an agent to autonomously place itself within a group.

## KEYWORDS

F-formations; Social Robots; Human-Robot Interaction

### ACM Reference Format:

Sai Krishna Pathi. 2018. Join the Group Formations using Social Cues in Social Robots. In *Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018), Stockholm, Sweden, July 10–15, 2018*, IFAAMAS, 2 pages.

## 1 INTRODUCTION

In examples of agent interaction such as Human-Robot Interaction (HRI), an agent is either operated manually or automatically should respect other agent's space, understand their behaviour, dynamics and intention behind their actions. For this, an agent needs to understand social signals, which include non-verbal behavioural cues such as facial expressions, body postures, gestures and proxemics. Proxemics in particular are very important in social interactions. The concept was developed by E.T. Hall [2]. Proxemics is the study that explains how people perceive and use space while interacting with each other. Robots need to understand, learn and execute this proxemics while interacting with humans. It is divided into four different zones: Intimate space, Personal space, Social space and Public space. We are mainly concerned with Social space as HRI falls under this zone. In social interactions, humans have a tendency to organise themselves in spatial patterns, while interacting with each other. Regarding these spatial arrangements, one of the promising framework is Adam Kendon's Facing formations [4] famously known as F-formations.

These F-formations are very helpful in increasing the quality of interaction and further could be used to have a collaborative work between humans, robots and agents. In order to have a social

interaction between humans and robot, the robot should be enabled to automatically adhere to F-formations while joining the groups. For this, *Firstly*, an agent should detect the formations in which people are standing. *Secondly*, find a spot in the group and navigate itself into the formation to socially interact with people.

Solving this problem is very useful for a variety of devices including both social and telepresence robots. In our case, we are studying both the robots. The project is about developing methods to enable the robots to join the groups and understanding the effect on users.

## 2 RESEARCH PLAN

The research plan, simply put, is divided into 3 phases, which are:

- (1) Phase I: Find the importance of F-formations in robotics.
- (2) Phase II: Develop methods to detect F-formations with respect to robot.
- (3) Phase III: Develop an approach to navigate the robot into the formations.

### 2.1 Phase I

In social interactions, people organise themselves in F-formations for better interaction as proposed by Kendon [4]. Using these F-formations, a better quality of interaction was observed between human and robot while socially interacting [3]. When telepresence robots were placed according to F-formations, the older people and the teleoperator had a nice quality of interaction [5] but we are not sure, would the teleoperator follow F-formations while teleoperating the robot and place the robot in the formations.

For this, we studied the behaviour of teleoperators of mobile telepresence systems. The purpose of the work was to determine to what extent teleoperators adhere to spatial and orientational relationships known as F-formations, while remotely interacting with groups. To prove this, we have drawn 3 expectations. The expectations are:

[E1]: *People teleoperating a mobile robot will respect F-formations while joining a social interaction or group.*

[E2]: *Teleoperators will consume time to place themselves within a configuration that it takes to simply approach the group.*

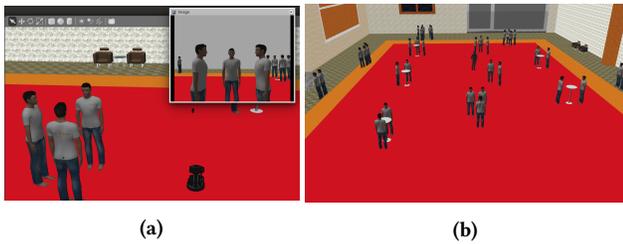
[E3]: *An autonomous feature is necessary to navigate the robot to join the social interaction or group.*

In order to validate our expectations, a simulated environment was created with simulated characters and conducted the experiment by inviting the participants. The simulation environment

---

**Main Supervisor: Prof. Amy Loutfi**, Head of AASS, Örebro University, Örebro, Sweden.

*Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018)*, M. Dastani, G. Sukthankar, E. André, S. Koenig (eds.), July 10–15, 2018, Stockholm, Sweden. © 2018 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). All rights reserved.



**Figure 1:** (a) *The robot observing the humans interacting and the inside image shows the view from robot's camera* (b) *Humans standing in different formations in the simulation environment.*

is a conference lobby where humans are having their break time after the sessions. From the Figure 1, we can observe the robot and the humans standing in different formations. The participants joined the groups through teleoperating a mobile robot. The evaluation was done using different tools, on one side a questionnaire was provided for participants to answer and on the other side a qualitative method was used to validate the formations made by the teleoperators. From the results obtained, we conclude that people teleoperating the mobile robot do respect F-formations while joining the groups or social interactions.

From this experiment, we have also found that teleoperators consume more time to place the robot in the group and there is a need for an autonomous feature to navigate the robot to join the social interactions.

In this phase, our virtual environment was used to verify that we needed a method to navigate the robot into the groups. The added advantage of virtual environment is that it allows for flexible and easily change experimental setup. In the coming phases, our work would involve working with a real robot and communicating with people using F-formations in a real time-scenario.

## 2.2 Phase II

From Phase I, we have found that there is a need for autonomous feature to join the groups. Developing an autonomous feature can be further divided into two phases, which are, Phase II, a method to detect F-formations and Phase III, an approach to navigate the robot into the groups or social interactions.

Regarding developing a method to detect F-formations, many researchers have proposed different methods to solve this problem. The recent state of the art algorithm is Zhang et al [8], where they consider the position and body orientation of people not only to detect F-formations but also the associates and singletons in the image. They propose a spatial-context-aware F-formation detector, which considers the influence of social and spatial context while modelling by learning the frustum of attention but their approach was from bird's view. The work which approaches this problem from egocentric view is [1], they estimate the head pose and 3D people locations to build a bird's view and a supervised clustering algorithm is used to detect groups. Most of the strategies are developed from computer vision and machine learning communities and none of them are implemented on robots.

The problem should be approached from a functional point of view such that the developed algorithm should work on a mobile robot with an egocentric camera, in real time, with low memory, less computational time and operate in natural scenarios. Vazquez et al [7] explored this problem in robotics community and proposed detecting the F-formations based on lower body estimation, which is obtained by tracking the position and orientation of the people in the scene but using an exocentric camera (overhead video data set). This method was evaluated on a 2D overhead video data set, which may not be easily scale to being integrated onto a mobile robot which does not have access to exocentric cameras. So, still there is an open question to detect F-formations in real time on a mobile robot in natural settings [6]. The developed methods suffer from one or more limitations of the following: considering prior information which are positional information (x,y), orientational information ( $\theta$ ) or both, does not work in real time and do not consider from robot's perspective (egocentric view).

For this, we did develop a preliminary approach to detect F-formations using Pepper robot. Further, we would propose a method, which would take into account the present limitations and handle the uncertainties, which would also detect multiple formations in one scene.

## 2.3 Phase III

In this phase, we would develop an algorithm to estimate the best spot in the group and navigate the robot automatically into the group to socially interact with people.

## 3 CONCLUSION

Our work is intended to build a software which could be integrated in mobile robots to have a smooth interaction between people and robot. In order for the robot to cope up with different situations, the robot should learn to manage the distance and orientation features based on the context.

## REFERENCES

- [1] Stefano Alletto, Giuseppe Serra, Simone Calderara, and Rita Cucchiara. 2015. Understanding social relationships in egocentric vision. *Pattern Recognition* 48, 12 (2015), 4082–4096.
- [2] Edward Twitchell Hall. 1966. *The hidden dimension*. (1966).
- [3] Wafa Johal, Alexis Jacq, Ana Paiva, and Pierre Dillenbourg. 2016. Child-robot spatial arrangement in a learning by teaching activity. In *Robot and Human Interactive Communication (RO-MAN), 2016 25th IEEE International Symposium on*. Ieee, 533–538.
- [4] Adam Kendon. 2010. Spacing and orientation in co-present interaction. In *Development of Multimodal Interfaces: Active Listening and Synchrony*. Springer, 1–15.
- [5] Annika Kristoffersson, Kerstin Severinson Eklundh, and Amy Loutfi. 2013. Measuring the quality of interaction in mobile robotic telepresence: A pilot's perspective. *International Journal of Social Robotics* 5, 1 (2013), 89–101.
- [6] Angélique Taylor and Laurel D Riek. 2016. Robot Perception of Human Groups in the Real World: State of the Art. In *AAAI Fall Symposium Series: Artificial Intelligence for Human-Robot Interaction Technical Report FS-16-01*. Retrieved January, Vol. 4. 2017.
- [7] Marynel Vázquez, Aaron Steinfeld, and Scott E Hudson. 2015. Parallel detection of conversational groups of free-standing people and tracking of their lower-body orientation. In *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*. IEEE, 3010–3017.
- [8] Lu Zhang and Hayley Hung. 2016. Beyond F-formations : Determining Social Involvement in Free Standing Conversing Groups from Static Images. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2016). <https://doi.org/10.1109/CVPR.2016.123>