29th IEEE International Conference on Robot & Human Interactive Communication ROMAN-2020
Crowdbot Workshop: Robots from Pathways to Crowds, ethical, legal and safety concerns of robots navigating human environments

Crowd Polarization as Environmental Alignment Heuristic

Ben Wright[1] Victoria Edwards[2]

***Abstract*— Social norms govern human society and are some of the most important rules that dictate interaction. Here, we focus on social norms for movement and path determination. To have robots safely interact in real world settings, it is critical that they follow these same norms. This requires a representation useful to the robot, allowing it to enter environments with increased understanding of how humans move in the space. In this work, we capture social movement norms by measuring the alignment of humans in an environment and we are able to see clear patterns emerge.**

I. Introduction & Background

Mapping and navigation through crowded environments has been a major focus for autonomous systems as robots leave the lab and enter the world. However, robots enter new environments blind to any social norms attributed to a location. For example, a robot that does not know that sidewalks are preferred over grass or not to navigate through a group playing soccer can be dangerous to both the robot and the individuals. This lack of contextual awareness puts both the robot and the humans in the environment at risk of interacting in a way that feels unnatural and can potentially be dangerous. Humans bring years of experience and context to new environments which allows them to integrate into settings easily. If a robot could have access to a representation of these social norms, it would make it easier for the robot to assimilate into the environment. We will use an alignment measurement to classify movement social norms into a map representation which a robot can use for navigation. Within this map we can plan for areas which may have more challenging behaviors the robot has to deal with, and also optimize paths such that the robot follows the standards of the people in the environment. We will discuss the possibility that a broader subset of social norms (e.g. walking on certain sides of the street, staying off of landscaping, etc.) can actually be transitioned from one environment to another.

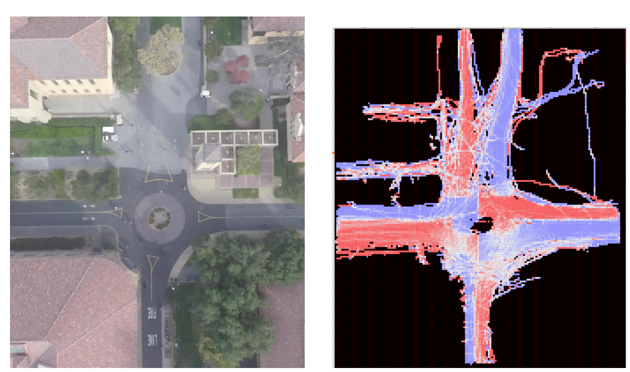
Swarming systems are characterized as agents that operate with simple underlying rules in a distributed and decentralized manner, which results in collective behavior. There are both natural and artificial swarms, and a common evaluation metric to use when studying collective behavior is to consider the swarm polarization [1]. Swarm polarization is a measure of alignment of all agents with respect to an origin or center of mass [2]. In this work, we will utilize data that observes humans following social movement norms. Under these movement norms there are many possible trajectories a human can take. We exploit the measure of alignment, from the swarming literature, to build a map representation of low level behavior of humans moving through the environment such that a robot can move through an unknown space consistent with the behaviors of people around it.

Within robotics, a variety of work has been done to allow robots to navigate smoothly through crowds of people including techniques using: flow maps [3], Reinforcement Learning (RL) [4], and Multi-Policy Decision Making (MPDM) [5]. A recent work addresses a robot navigating around a group of pedestrians to a goal using Deep RL and prediction of the intent of the people [6]. A specific example which relates to what we are proposing builds flow maps of observed human trajectories by robots and predicting future flow [7]. These approaches are all intended to be used to aid the robot while it is actively moving in the environment. Our approach is information that can be used in advance of the robot entering an environment to allow the robot to take paths and trajectories that are common and natural to other people using the same space.

II. Method

We used the Stanford Drone Dataset (SDD)[3] as a basis to explore social norms. The SDD consists of overhead videos from eight different locations around the Stanford University campus with a focus on people moving throughout the environment. The locations contain sidewalks, bike paths, roads, and many other types of environments. Object paths were extracted using computer vision techniques described fully in [8] and contain pedestrians, bikes, golf carts, skateboarders, and buses. Each object was given a class (e.g., pedestrian) and a bounding box for each frame of the video. The SDD contains over 19K objects and consists of many different

[3] `http://cvgl.stanford.edu/projects/uav_data/`



(a) View of Overhead Environment (b) Generated Alignment Map

Fig. 1. A visualization of the polarization calculated for a given area.

[1] NRC Postdoctoral Associate, US Naval Research Laboratory
[2] US Naval Research Laboratory

types of interactions, groups, and object motion. Using these annotations, we previously [9] adapted a number of extra annotations, most notably we collapsed the bounding box to its midpoint (x,y). Each object, $\mathcal{O}_j$, has a location per frame, $\mathcal{O}^i_j$. With these annotations we were able to calculate a measure of alignment from the SDD data. First, we needed to understand what we are measuring our alignment with respect too, this could be thought of as the origin in a standard Euclidean graph. By considering alignment around a central point we are able to consider many moving objects and build an average representation of alignment instead of observing local trajectories. However, because we are working in pixel space, we hand selected the origins in the videos to be the middle of intersections or roundabouts. We will refer to this hand selected origin as a landmark, $\mathcal{L}$. An example landmark is the center roundabout in Figure 1.

To compute alignment of objects in each frame, we modify the swarm polarization which typically computes the normalized average alignment for all agents in the swarm [1]. The modified polarization is now a measure of alignment to capture social norms. This value is computed for each object and stored at the position of the object, written as follows:

$$\mathcal{A}_{\mathcal{O}_j} = \frac{(\sum_i^N \Delta d_i(\mathcal{O}_j)^2}{\sum_i^N \Delta d_i(\mathcal{O}_j)^2}, \quad (1)$$

where N is all frames in the data set, and $\Delta d_i(\mathcal{O}_j)$ is the change in distance between an object, $\mathcal{O}_j$, and a landmark from frame to frame. For spaces with no objects the value in the alignment map is NaN. This results in positive values for objects that go farther away from the landmark and negative values for objects that get closer. This distance value is computed as follows:

$$\Delta d_i(\mathcal{O}_j) = ||\mathcal{L} - \mathcal{O}^i_j|| - ||\mathcal{L} - \mathcal{O}^{i+1}_j||, \quad (2)$$

where $\mathcal{O}^i_j$ and $\mathcal{O}^{i+1}_j$ is the position of the object from frame f_i to frame f_{i+1} within the annotated data. This can be done for all objects in each frame.

We do not necessarily need, nor are interested in, the alignment at the pixel level, because not every space is occupied (e.g. some areas are buildings where no people walk over buildings). In order to separate out the alignment of the whole video, we used sections of the video in $10x10$ pixel squares, $\mathcal{G}$. So, for each $\mathcal{G}$, we get the average alignment for all objects within that grid. We do this irrespective of time or duration within grid, such that:

$$\mathcal{A}_{\mathcal{G}} = \frac{1}{P_{\mathcal{G}}} \sum_j^{P_{\mathcal{G}}} \mathcal{A}_j, \quad (3)$$

where P is the number of objects, $\mathcal{O}$, in the specific $\mathcal{G}$. These equations were encoded and run across the SDD data for preliminary results in building social norm maps for robots.

III. Preliminary Results

Using equation 3 alignment maps were created for the SDD. The landmark used for alignment was the center of the roundabout as seen in Figure 1(a). The resulting alignment map can be seen in Figure 1(b). From the alignment calculations, there are 4 values of interest: A value of -1 indicates a strong alignment *towards* the landmark; A value of 1 indicates a strong alignment *away* from the landmark; A value of 0 indicates no alignment pattern; and a non-existent value indicates no presence of moving objects. These 4 values are shown by the colors red, blue, white, and black respectively.

Figure 1(a) shows a simple environment of a roundabout with walking paths around it. We can see patterns of strong alignment emerge for this environment in Figure 1(b), not only around the roundabout but along walking paths as well. Very few areas appear to have a weak-alignment, which tend towards the color white.

IV. Discussion

This work encodes information that humans have into something that is usable by a robot to allow for safe and natural integration of the robot into a human environment. Standard navigation, mapping, and obstacle avoidance for a robot do not often consider the rules with which humans interact in an environment, instead they focus on real time avoidance or mapping of what people are doing. By encapsulating the social movement norms that are attributed to an environment we hypothesize that we can have a robot take these movement norms of others around it in order to act more safely (i.e., walk on the right side of the sidewalk rather than the left or middle, as suggested in figure and results above).

References

[1] V. Edwards, P. deZonia, M. A. Hsieh, J. Hindes, I. Triandaf, and I. B. Schwartz, "Delay Induced Swarm Pattern Bifurcations in Mixed Reality Experiments," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 30, no. 7, p. 073126, 2020. [Online]. Available: https://doi.org/10.1063/1.5142849

[2] K. Tunstrøm, Y. Katz, C. C. Ioannou, C. Huepe, M. J. Lutz, and I. D. Couzin, "Collective States, Multistability and Transitional Behavior in Schooling Fish," in *PLoS Computational Biology*, 2013.

[3] R. Bresson, J. Saraydaryan, J. Dugdale, and A. Spalanzani, "Socially Compliant Navigation in Dense Crowds," in *2019 IEEE Intelligent Vehicles Symposium (IV)*, 2019, pp. 64–69.

[4] P. Henry, C. Vollmer, B. Ferris, and D. Fox, "Learning to Navigate Through Crowded Environments," in *2010 IEEE Intl. Conf. on Robotics and Automation*, 2010, pp. 981–986.

[5] A. Cunningham, E. Galceran, D. Mehta, G. Ferrer, R. Eustice, and E. Olson, *MPDM: Multi-policy Decision-Making from Autonomous Driving to Social Robot Navigation*, January 2019, pp. 201–223.

[6] A. J. Sathyamoorthy, U. Patel, T. Guan, and D. Manocha, "Frozone: Freezing-Free, Pedestrian-Friendly Navigation in Human Crowds," *IEEE Robotics and Automation Letters*, vol. 5, no. 3, pp. 4352–4359, 2020.

[7] F. Jumel, J. Saraydaryan, and O. Simonin, "Mapping Likelihood of Encountering Humans: Application to Path Planning in Crowded Environment," in *2017 European Conf. on Mobile Robots (ECMR)*, 2017, pp. 1–7.

[8] A. Robicquet, A. Sadeghian, A. Alahi, and S. Savarese, "Learning Social Etiquette: Human Trajectory Understanding in Crowded Scenes," in *European Con. on Computer Vision*. Springer, 2016, pp. 549–565.

[9] B. Wright, M. Bugajska, W. Adams, E. Lawson, J. M. McCurry, and J. G. Trafton, "Proxemic Reasoning for Group Approach," in *12th Intl. Conf. on Social Robotics (ICSR)*, Under Review.