CoMet: Modeling Group Cohesion for Socially Compliant Robot Navigation in Crowded Scenes

Adarsh Jagan Sathyamoorthy, Utsav Patel, Moumita Paul, Nithish K Sanjeev Kumar, Yash Savle, and Dinesh Manocha

***Abstract*— We present CoMet, a novel approach for computing a group's cohesion and using that to improve a robot's navigation in crowded scenes. Our approach uses a novel cohesion-metric that builds on prior work in social psychology. We compute this metric by utilizing various visual features of pedestrians from an RGB-D camera on-board a robot. Specifically, we detect characteristics corresponding to proximity between people, their relative walking speeds, the group size, and interactions between group members. We use our cohesion-metric to design and improve a navigation scheme that accounts for different levels of group cohesion while a robot moves through a crowd. We evaluate the precision and recall of our cohesion-metric based on perceptual evaluations. We highlight the performance of our social navigation algorithm on a Turtlebot robot and demonstrate its benefits in terms of multiple metrics: freezing rate (57% decrease), low deviation (35.7% decrease), path length of the trajectory(23.2% decrease).**

I. Introduction

Mobile robots are increasingly used in crowded scenarios in indoor and outdoor environments. The underlying applications include surveillance, delivery, logistics, etc. In such scenarios, the robots need to navigate in an unobtrusive manner and also avoid issues related to sudden turns or freezing [1]. Moreover, the robots need to integrate well with the physical and social environment.

Extensive research in social and behavior psychology suggests that that crowds in real-world scenarios are composed of (social) groups. A group is generally regarded as a meso-level concept and corresponds to two or more pedestrians that have similar goals over a short or long periods of time. As a result, the pedestrians or agents in that group exhibit similar movements or behaviors. It is estimated that up to 70% of observed pedestrians in real-world crowds are part of a group [2], [3]. Therefore, it is important to understand group characteristics and dynamics and use them to perform socially-compliant robot navigation [4], [5], [6].

The problem of efficient robot navigation among pedestrians has been an active area of research. Most existing robot navigation algorithms consider walking humans or pedestrians as separate obstacles [7], [4], [6], [8]. Some techniques tend to predict trajectories of each pedestrian using learning-based methods [8], but do not account for the influence of group characteristics on individuals. This could lead to obtrusive trajectories that may cut through groups of friends or families. Other methods use simple and conservative methods to detect locally sensed clusters of pedestrians and compute path around them [6]. However, they do not work well as the crowd density increases.

One characteristic of groups that could be utilized to address these problems is the social-cohesion or the collective behavior of its members. This is directly linked to the interpersonal relationship between group members. For example, a group of friends or family has high cohesion, as compared

(a) (b)

Fig. 1: Our novel navigation method uses our CoMet metric to compute a collision-free trajectory for a robot in real-world scenarios. CoMet identifies groups in crowds and detects intra-group proximity, walking speed, group size and interactions to estimate a group's cohesion. (a) In dense scenarios, our navigation algorithm identifies a low-cohesion group (red bounding box) and navigates between its group members (green path) by assuming human cooperation for navigation; (b) Our method detects a high cohesion group (green bounding box) and plans a trajectory around it. Overall, our method improves social-compliance and the naturalness of the trajectory (Section V-A).

to a group of strangers [9], [10], and inversely related to the *permeability* of the group in social settings, i.e. whether another individual can cut through the group while walking [11]. Many theories have been proposed in psychology and sociology to identify the human behaviors or features that are regarded as good indicators of group cohesion. Such features include proximity between group members [12], walking speed [13], group size [11], context or environment [14], etc. Estimating cohesion could help a robot plan a better or socially compliant trajectory based on the context. For example, in dense crowds (i.e. pedestrian density is more than 1 person/m^2), the robot could navigate around a group that has high cohesion. Or the robot could move between members of a group that has low cohesion, similar to how humans navigate in crowded scenarios.

Main Contributions: We present a novel algorithm to perform socially compliant navigation in crowded scenes. Our approach uses perception algorithms to identify groups in a crowd using visual features. We also present a novel group cohesion metric and efficient algorithms to compute this metric in arbitrary crowds using deep learning. We combine our cohesion metric with learning-based techniques to generate trajectories that tend to follow the social norms. Some of the novel components of our approach include:

- We present CoMet, a novel metric for estimating group

cohesion. Our approach is based on social psychology studies and exploits features such as proximity between people, walking speeds, group sizes, and interactions. Our method uses an RGB-D sensor to detect groups and these visual features. CoMet has a near 100% precision and recall when identifying low-cohesion groups, when evaluated in real-world pedestrian or crowd datasets.
- We present a novel CoMet-based navigation method that accounts for group cohesion, while ensuring social-compliance in terms of naturalness, large deviations and freezing behaviors. Our formulation assumes human cooperation in dense crowds and plans less conservative trajectories, as compared to prior methods. We prove that the deviation angles computed by our method are less than or equal to the deviation angles computed using a prior social navigation algorithm [6].
- We implement CoMet on a real Turtlebot robot equipped with a commodity RGB-D sensor and demonstrate improvements in terms of social navigation. Our qualitative evaluations in dense scenes indicate that CoMet accurately identifies the cohesion in different groups (see Fig.1). This enables the Turtlebot robot to navigate through a group based on our cohesion metric. As compared to prior social navigation algorithms, we demonstrate improved performance in terms of following metrics: freezing rate (up to 57% decrease), path deviation or turns (up to 35.7% decrease), and path length (up to 23.2% decrease).

II. Related Works

In this section, we briefly review prior work in robot navigation among crowds, pedestrian and group detection, and group interactions.

A. Clustering-based Group Detection

Many techniques have been proposed in computer vision for pedestrian and group detection in a crowd. The first step in group detection is to detect individuals in the images or videos. These include many deep learning-based methods for pedestrian detection and tracking [15] and improved methods for high density crowds [16]. These methods have been extended from individual pedestrian detection to group detection [17], [18]. These group-based methods typically use different kinds of clustering based on the proximities between people, their trajectories and their velocities to segregate them into groups [18], [19].

B. Group Behavior, and Interaction Detection

Different techniques have been proposed for detecting group behaviors and interactions in computer vision and social psychology [14], [20] as well as event identification [21]. Behavior detection and event identification involve the analysis of different features (e.g., collectiveness, stability, uniformity) that represent how people move and interact in a crowd. These include individuals, groups, leaders, followers etc. They also involve detecting scenarios where groups merge together or split while walking or running.

Another set of relevant techniques involve detecting interacting among people in a group based on F-formations [17], [22]. These algorithms estimate features such as people's body and head poses and identify the individuals who are facing each other. Our approach is complimentary to these methods and extends them by using many other features, inculding proximity, walking, interaction, to gauge group cohesion.

C. Robot Navigation and Social Compliance

Many recent works have focused on socially-compliant navigation [23], [4], [5], [6], [24], [25]. The underlying goal is to design methods that not only compute collision-free trajectories but also comply with social norms that increase the comfort level of pedestrians in a crowd. At a broad level, the three major objectives of social navigation are comfort, naturalness, and high-level societal rules [23]. For example, a robot needs to avoid movements that are regarded as obtrusive to pedestrians by following rules related to how to approach and pass a pedestrian [4], [6]. Other techniques are based on modeling intra-group interactions [24] or by learning from real-world static and dynamic obstacle behaviors [26].

Many techniques for social navigation have been proposed based on reinforcement learning (RL) or inverse reinforcement learning (IRL). The RL-based methods [7], [27], [28], [6] mostly focus on treating each pedestrian as a separate obstacle, avoid collisions or sudden turns or large deviations. IRL methods are driven by real-world natural crowd navigation behaviors [29], [30] and used to generate trajectories with high level of naturalness. However, they can result in unsafe trajectories and may not work well as the crowd density increases. Some methods model pedestrian behaviors by learning about their discrete decisions and the variances in their trajectories [31]. Other works have modeled human personality traits [25] or pedestrian dominance [32] based on psychological characteristics for trajectory prediction and improved navigation. Bera et al. [33] present an algorithm to avoid negative human reactions to robots by reducing the entitativity of robots. Our work on modeling group cohesion is complimentary to these methods.

III. Background

In this section, we give an overview of prior work in social psychology, pedestrian tracking and robot navigation that are used in our approach. We also introduce the symbols and notation used in the paper.

A. Social Psychology

We use four features to estimate the cohesion of a group based on prior work in social psychology. We give a brief overview of each of these features.

Proximity: Proximity is chosen based on the proxemics principles established by Hall [12]. The underlying theory states that humans have an intimate space, a social and consultative space and a public space when interacting with others. We extend this idea to unstructured social scenarios where people or pedestrian walk, stand or sit together. In general, humans maintain a closer proximity to other people with whom they closely interact (high cohesion).

Walking Speed: [13] studied individual and mixed gender groups' walking speeds in a controlled environment and observed significantly slower speeds when people walk with their romantic partners. *Assertion 1:* A slower-than-average walking speed in a group indicates a close relationship between group members (high cohesion)

Group Size: [11] analyzed the perceptions of people when passing through a group of two and four people in a university hallway. It was observed that people tend to

Symbols	Definitions
$\mathbf{p}^{i,t}$	Position vector of person i at time t relative to the robot/camera coordinate frame. $\mathbf{p}^{i,t} = [x^{i,t}, y^{i,t}]$
$\mathbf{v}^{i,t}$	Walking velocity vector of person i at time t relative to the robot/camera coordinate frame. $\mathbf{v}^{i,t} = [\dot{x}^{i,t}, \dot{y}^{i,t}]$
I^t_{rgb}	RGB image captured at time instant t of width w and height h.
I^t_{depth}	Depth image captured at time instant t of width w and height h.
$\mathbb{B}^i$	Bounding box of person with ID i.
$[x^{\mathbb{B}^i}_{cen}, y^{\mathbb{B}^i}_{cen}]$	Centroid of the bounding box $\mathbb{B}^i$.
$\mathbf{x}^{i,t}$	State vector used in Kalman filter for walking vector estimation for person i at time instant t. $\mathbf{x}^{i,t} = [x^{i,t}\ y^{i,t}\ \dot{x}^{i,t}\ \dot{y}^{i,t}]^T$
ID^t	Set of person IDs detected in I^t_{rgb}.
n^k	Number of members in Group $G^{k,t}$.
N	Number of people detected in an RGB image.
$\mathbf{f}^i_p, \mathbf{f}^i_o$	Vectors for person i's face position and orientation relative to the camera coordinate frame.
K_p, K_w, K_s K_i	Proportionality and weighing constants for each feature
$CH()$	Convex Hull function with points as its arguments.

TABLE I: List of symbols used in CoMet and their definitions.

penetrate through the 4-person group less than the 2-person group. *Assertion 2:* Humans perceive the cohesion of a bigger group to be higher (implying low permeability) than that of a smaller group. Permeability of a group is a measure of the resistance that a moving non-group entity faces while passing in-between the members of a group.

Interactions: Jointly Focused Interactions (JFI) [34] entail a sense of mutual activity and engagement between people and imply their willingness to focus their attention on others. Therefore, it is chosen as an indicator for cohesion. We extrapolate JFI to signify a higher level of cohesion between group members. Visually, interactions can be detected by estimating peoples' 3-D face vectors [22] and detecting when the vectors point towards each other.

B. Notations and Definitions

We highlight the symbols and notation used, in Table I. We use *i*, *j*, and *k* to represent indices. All distances, angles and velocities are measured relative to a rigid coordinate frame attached to the camera (on the robot) used to capture the scene. The X-axis of this frame points outward from the camera and the Y-axis points to the left, with its origin at the center of the image. The time interval between two consecutive RGB images in the stream is Δt.

C. Frozone

Frozone [6] is a navigation method that tackles the Freezing Robot Problem (FRP) [1] arising in crowds. At the same time, it can generate trajectories that are less obtrusive to pedestrians. The underlying algorithm computes Potential Freezing Zone (PFZ), which corresponds to a configuration of obstacles where the robot's planner halts the robot and starts oscillating for a period of time when it deems that all velocities could lead to a collision. Frozone's formulation is conservative and not suitable for dense crowd navigation ($>$ 1 person/m^2). In addition, it treats pedestrians as individual obstacles. The main steps in Frozone's formulation are: (1). identifying *potentially freezing* pedestrians who could cause FRP based on their walking vectors and predicting their positions after a time horizon t_h (2). constructing a Potential Freezing Zone (PFZ_{froz}) as the convex hull of the predicted positions $\hat{\mathbf{p}}^{i,t+t_h}_{pf}$ (see 5-sided convex polygon in Fig. 3) for all the *potentially freezing* pedestrians. PFZ_{froz} is formulated as,

$$PFZ = ConvexHull(\hat{\mathbf{p}}^{i,t+t_h}_{pf}), \quad i \in 1, 2, ..., P_f. \tag{1}$$

and (3). computing a deviation angle for the robot to avoid PFZ_{froz} if its current trajectory intersects with it. PFZ_{froz} corresponds to the set of locations where the robot has the maximum probability of freezing and being obstructive to the pedestrians around it. The deviation angle to avoid it is computed as

$$\phi_{froz} = min(\phi_1, \phi_2), \tag{2}$$

where ϕ_1 and ϕ_2 are given by,

$$\phi_1 = \underset{R_{z,\phi_1}\mathbf{v}_{rob}t_h \notin PFZ_{froz}}{\operatorname{argmin}} \left(dist(R_{z,\phi_1} \cdot \mathbf{v}_{rob} \cdot t_h,\ \mathbf{g}_{rob})\right), \tag{3}$$

$$\phi_2 = \tan^{-1}(y^{near,t}/x^{near,t}), \quad \phi_2 \neq 0. \tag{4}$$

Here, R_{z,ϕ_1} is the 3-D rotation matrix about the Z-axis (perpendicular to the plane of the robot), $\mathbf{v}^{rob}, \mathbf{g}^{rob}$ represent the current velocity and the goal of the robot. $[x^{near,t}, y^{near,t}]$ denotes the current location of the nearest freezing pedestrian relative to the robot. This point is in the PFZ's exterior. For navigating the robot towards its goal, along with handling static obstacles and dense crowds, a Deep Reinforcement Learning (DRL)-based method [28] is used. However, the resulting navigation may cut through groups regardless of their cohesion (see Fig. 1(a)). As a result, the robot's trajectory may not be socially compliant.

IV. CoMet: Modeling Group Cohesion

In this section, we present our group cohesion metric, which first classifies pedestrians into groups and then measures their closeness or cohesion. Our method runs in real-time, taking a continuous stream of RGB and depth images as input and detecting the group features highlighted in Section III-A. Our overall approach based on these features is shown in Fig. 2.

A. Detecting Group Features

In this section, we first describe how we track and localize people, detail conditions for a set of people to be classified as a group and then explain efficient techniques to detect group features from RGB and depth images.

B. Pedestrian Tracking and Localization

A key issue in detecting the features mentioned in Section III-A is to first detect, track and localize each pedestrian position relative to the camera frame in a continuous stream of RGB images. We use YOLOv5 [35] and Deep Sort [36] algorithms to detect people and track people, respectively. YOLOv5 outputs a set of bounding boxes $\mathcal{B} = \{\mathbb{B}^i\}$ for each detected pedestrian *i* in an RGB image at time instant t (denoted as I^t_{rgb}). $\mathbb{B}^i$ is denoted using its top-left and bottom-right corners in the image-space or pixel coordinates. In addition, we also assign a unique integer number as an ID for each detected pedestrian.

Next, to accurately localize people, the distance of each detected person relative to the camera coordinate frame must be estimated. To this end, we use a depth image I^t_{depth}, whose every pixel of which contains the proximity (in meters) of an object at that location of the image. The pixels

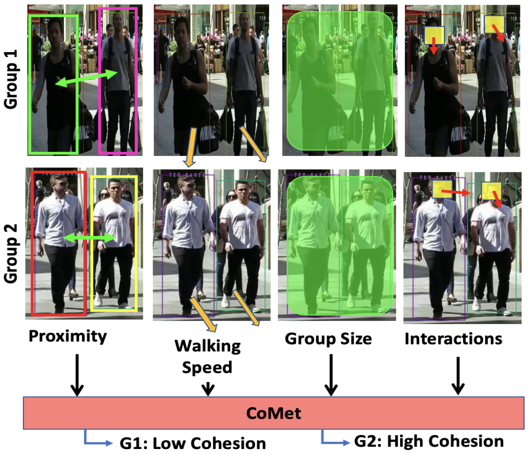

Fig. 2: Computation of Group Cohesion Metric: We use four features to compute the metric. We highlight how our approach is used on two different groups of pedestrians at a given time instant. These groups are shown in different rows. Group 1 consists of two pedestrians in close proximity, but they are walking away from each other. Their distance is increasing in subsequent frames and they are looking in different directions. This implies little interaction and low group cohesion. Group 2 has two people walking together with their faces are oriented towards each other, which indicates high interaction and a high group cohesion.

in I^t_{depth} contain values between a minimum and maximum distance range, which depends on the camera used to capture the image.

1) Group Classification: Let us consider any set of people's IDs $G^{k,t} \subseteq ID^t$. At any time t, if the following conditions hold,

$$\|\mathbf{p}^{i,t} - \mathbf{p}^{j,t}\|_2 \leq \Gamma \quad \forall i,j \in G^{k,t} \tag{5}$$

$$\|\mathbf{p}^{i,t} - \mathbf{p}^{j,t}\|_2 \geq \|\mathbf{p}^{i,t} + \mathbf{v}^{i,t} - (\mathbf{p}^{j,t} + \mathbf{v}^{j,t})\|_2. \tag{6}$$

then the set $G^{k,t}$ is classified as a *group* in the image I^t_{rgb}. Here, Γ is a distance threshold set manually and $\mathrm{sign}()$ is the signum function. The first condition ensures that people are close to each other and the second condition ensures that the group members walk in the same direction. When $|\mathbf{v}^i\|, \|\mathbf{v}^j\| = 0$ (static groups), only the first condition is used for grouping.

2) Estimating Proximity: To estimate the proximity between people at time instant t, first the bounding boxes detected in the RGB image by YOLOv5 are superimposed over the depth image. To estimate the distance of a person i from the camera ($d^{i,t}$), the mean of all the pixel values within a small square centered around $[x^{\mathbb{B}^i,t}_{cen}, y^{\mathbb{B}^i,t}_{cen}]$ is computed. The angular displacement $\psi^{i,t}$ of person i relative to the camera can be computed as, $\psi^{i,t} = \left(\frac{\frac{w}{2} - x^{\mathbb{B}^i,t}_{cen}}{w}\right) * FOV_{RGBD}$. Here FOV_{RGBD} is the field of view of the RGB-D camera. Person i's location relative to the camera can be computed as $[x^{i,t}\ y^{i,t}] = d^{i,t} * [\cos\psi^{i,t}\ \ \sin\psi^{i,t}]$. The distance between a pair of people i and j can then be computed as, $dist(i,j) = \sqrt{(x^{i,t} - x^{j,t})^2 + (y^{i,t} - y^{j,t})^2}$.

3) Estimating Walking Speed and Direction: To estimate the i^{th} person's walking vector $\mathbf{v}^i$ in I^t_{rgb}, we use a Kalman filter with a constant velocity motion model. All the detected people in I^t_{rgb} (with their IDs stored in the set ID^t) are modeled using the state vector $\mathbf{x}^t$ defined in Table I. If ID^t contains IDs which were not present in $ID^{t-\Delta t}$, we initialize their corresponding state vectors $\mathbf{x}^t$ with constant values. For all the pedestrians who were detected in previous RGB images, i.e., with previously initialized states, we update their states using the standard Kalman prediction and update steps [37]. We use a zero mean Gaussian noise with a pre-set variance to model the process and sensing noise.

4) Estimating Group Size: The size of the group can be trivially computed as the number of IDs in the set $G^{k,t}$.

5) Detecting Interactions: We use two 3-D vectors to represent the position and orientation of a person's face in I^t_{rgb}. We use OpenFace [38] to localize person i's face position relative to the camera coordinate frame ($\mathbf{f}^i_p$) on an RGB image and to obtain a unit vector ($\mathbf{f}^i_o$) for the face orientation. Two individuals are considered to be interacting if their face positions and orientations satisfy the following condition:

$$\|\mathbf{f}^i_p - \mathbf{f}^j_p\|_2 > \|\mathbf{f}^i_p + \mathbf{f}^i_o - (\mathbf{f}^j_p + \mathbf{f}^j_o)\|_2. \tag{7}$$

This condition checks if the distance between two people's face positions is greater than the distance between the points computed by extrapolating the face orientations (see dashed lines in Fig. 1). That is, if they are facing each other. We reasonably assume that non-interacting people do not face each other.

C. Cohesion Components

We now discuss how the detected group features can be used for cohesion estimation. Using multiple features to compute cohesion is advantageous in scenarios where not all features are properly senseable.

1) Proximity Cohesion Score: We use the average distance between group members to model the Hall's proxemics theory as previously extrapolated. As observed in Section III-A, cohesion between people is inversely proportional to the distance between them. Therefore, the cohesion score due to proximity is formulated as the reciprocal of the mean distance between group members as,

$$C_p(G^{k,t}) = K_p \cdot \frac{n^k}{\sum\limits_{\substack{i,j \in G^{k,t} \\ i \neq j}} dist(i,j)}. \tag{8}$$

2) Walking Speed Cohesion Score: Based on the discussion in Section III-A, we next compare the average walking speeds of each group with the average walking speed of all the detected people in I^t_{rgb}. Therefore, the cohesion score for a walking group ($\|\mathbf{v}^j\| \neq 0\ \forall j \in G^{k,t}$) due to its walking speed is formulated as,

$$C_w(G^{k,t}) = K_w \cdot \left(\frac{\sum\limits_{\forall i \in ID^t} \|\mathbf{v}^i\|}{N}\right) \Big/ \left(\frac{\sum\limits_{\forall j \in G^{k,t}} \|\mathbf{v}^j\|}{n^k}\right). \tag{9}$$

This reflects assertion 1 made in Section III-A, since cohesion is inversely proportional to walking speed. The average walking speed of the entire scene is included in this formulation to normalize out the effects of crowding in the scene. If $\sum\limits_{\forall j \in G^{k,t}} \|\mathbf{v}^j\| = 0$, i.e., the group is static, then $C_w(G^{k,t}) = K_w \cdot \eta$, where η is a user-set large constant value.

3) Group Size Cohesion Score: Based on assertion 2 in Section III-A, the cohesion of a group k is directly proportional to the group size (n^k). Therefore, the group size cohesion score is computed as,

$$C_s(G^{k,t}) = K_s \cdot n^k. \tag{10}$$

4) Interaction Cohesion Score: The interaction condition between any two people in a group (Equation 7) can be applied to all pairs in a group, and it's contribution to the cohesion score of a group can be re-written as:

$$C_i(G^{k,t}) = K_i \cdot \frac{1}{n^k} \cdot \sum_{\substack{i \neq j \\ i,j \in G^{k,t}}} \frac{\text{sign}(\theta_{ij})}{\cos\theta_{ij}} \quad \theta_{i,j} \in [-\frac{\pi}{4}, \frac{\pi}{4}]. \tag{11}$$

Here θ_{ij} is the angle between face orientation vectors $\mathbf{f}_o^i$ and $\mathbf{f}_o^j$ in the X-Y plane of the camera coordinate system. θ_{ij} is limited to $[-\frac{\pi}{4}, \frac{\pi}{4}]$ since face orientations are accurate in this range. Intuitively, we want the cohesion score to be positive and greater than 1, when people are facing towards each other and negative otherwise. Therefore, we choose the ratio $\frac{\text{sign}(\theta)}{\cos\theta}$, as it belongs to the range $[-\sqrt{2}, -1) \cup \{0\} \cup [1, \sqrt{2}]$ when $\theta_{i,j} \in [-\frac{\pi}{4}, \frac{\pi}{4}]$. Since $\cos()$ is an even function, the $\text{sign}()$ function ensures that the formulation is sensitive to the sign of the angle.

D. CoMet: Overall Group Cohesion Metric

Using the individual cohesion scores in Equations 8, 9, 10,11, the total cohesion score for a group at time t ($G^{k,t}$) is given as,

$$C_{tot}(G^{k,t}) = (C_p + C_w + C_s + C_i). \tag{12}$$

Here, $G^{k,t}$ is omitted in the RHS for readability. Note that K_p, K_w, K_s, K_i weigh the different features before adding them to the total cohesion score. If any of the features are not detectable, their contribution to C_{tot} will be zero. This acts as a measure of confidence, as our approach is able to better compute a group's cohesion when more features are detected.

Proposition IV.1. *The value of the overall cohesion metric $C_{tot}(G^{k,t})$ for a group is bounded.*

Proof. The proof to the proposition follows from the fact that C_p, C_w, C_s, C_i are bounded. The value of $C_p \in (0, K_p \cdot \Gamma]$, since Γ is used as a threshold to group people. The maximum value of C_w is $K_w \cdot \eta$, a large finite constant that is used when a group is static. C_s is bounded above by $K_g \cdot n^k$, which is finite. $C_i \in [-\sqrt{2}K_i, -K_i) \cup [K_i, \sqrt{2}K_i]$. ■

We use these bounds on the cohesion metric to compute appropriate thresholds that are used to categorize groups as low, medium and high cohesion groups.

V. Cohesion-based Navigation

In this section, we present our socially-compliant navigation algorithm that uses the group cohesion metric.

A. Socially-Compliant Navigation

Our objective is to improve the naturalness of a robot's trajectory. We attribute three quantities to trajectories with naturalness: 1. not suddenly halting or freezing (avoiding FRP), 2. low deviations angles, 3. not cutting in-between high cohesion groups (friends, families etc) in a crowd. This is in accordance with humans' walking behaviors where people do not suddenly halt or take large deviations from their goals [39], and do not cut through high cohesion groups while walking [11].

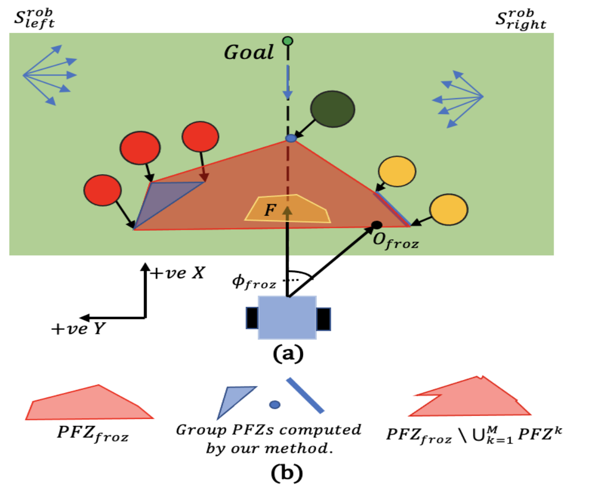

Fig. 3: **(a)** This scenario shows two groups (red and yellow agents in close proximity) and an individual pedestrian (green) walking towards the robot (blue box). The green rectangle denotes the robot's sensing region and the blue arrows denote the potentially freezing walking directions within each half of the sensing region. **(b) [Left]** PFZ_{froz} (a large 5-sided convex polygon in red) computed by Frozone [6] while considering each individual as a separate obstacle. **[Middle]** Our CoMet-based approach identifies each group and computes the group PFZs (PFZ^k). This corresponds to the blue triangle for the agents in the red group, a line for the agents in the yellow group, a point for the green individual. **[Right]** The region that represents $PFZ_{froz} \setminus \bigcup_{k=1}^{M} PFZ^k$. These shapes are shown separately to observe the differences. Our proposed method is less conservative and results in a smaller deviation from PFZs (no deviation needed in this case) than Frozone [6]. We also highlight one possible subset of $PFZ_{froz} \setminus \bigcup_{k=1}^{M} PFZ^k$, which contains positions with deviations $\leq \phi_{froz}$.

We extend Frozone [6] (Section III-C) by considering groups and their cohesions, and prove that our proposed method leads to smaller deviations from the robot's goal, and shorter trajectory lengths. It also does not navigate the robot through high cohesion groups. We assume a higher density in the environment (in terms of crowds and static obstacles) than Frozone's formulation and human cooperation for the robot's navigation. Frozone prevents the robot from moving in front of a pedestrian to avoid slow down in terms of their walking speeds.

B. CoMet-Based Navigation

To improve the naturalness of a robot's trajectory, our proposed method includes the following steps: (1). identifying potentially freezing groups within the sensing region of the robot and predicting their positions after a time period t_h; (2). constructing a PFZ for each group using the predicted future locations of each group member (see blue regions in Fig. 3); (3). computing a deviation angle to avoid the group PFZs while accounting for every group's cohesion. If a feasible solution is not found, the robot navigates in-between the group with the lowest cohesion in the scene.

Definition V.1 (*Potentially Freezing Groups:*) Groups of pedestrians that have a high probability of causing FRP after time t_h. Such groups are identified based on conditions of their average walking direction [6] (see blue arrows in Fig. 3). Groups that satisfy these conditions move closer to the robot as time progresses (proven in [6]). We predict the future positions of the potentially freezing group members as

$$\hat{\mathbf{p}}_{pf}^{i,t+t_h} = \mathbf{p}_{pf}^{i,t} + \mathbf{v}_{avg}^{G^{k,t}} t_h, \quad i \in G^{k,t}, \quad k \in \{1, 2, ..., M.\}. \tag{13}$$

Here, $\mathbf{v}_{avg}^{G^{k,t}}$ is the average group walking vector, computed as the mean of the walking vectors of the group members and, M is the total number of potentially freezing groups.

Definition V.2 (*Group PFZ*) The region in the vicinity of a group where the robot has a high probability of freezing. Instead of constructing the single PFZ as the convex hull of all potentially freezing pedestrians (like in [6]), we construct a PFZ for each potentially freezing group (see Fig. 3) as

$$PFZ^k = CH(\hat{\mathbf{p}}_{pf}^{i,t+t_h}), \; i \in G^{k,t}, \; k \in \{1, 2, ..., M.\}. \tag{14}$$

Every potentially freezing group's PFZ is computed for a future time instant $t + t_h$.

1) Computing Deviation Angle: If the robot's current trajectory navigates it into any of the group PFZs (implying an occurrence of FRP after time t_h), a deviation angle ϕ_{com} to avoid it is computed. The robot's current velocity $\mathbf{v}_{rob}$ is deviated by ϕ_{com} using a rotation matrix about the Z-axis as,

$$\mathbf{v'}_{rob} = R_{z,\phi_{com}} \cdot \mathbf{v}_{rob}, \tag{15}$$

$$\phi_{com} = \underset{\mathbf{v'}_{rob} \cdot t_h \notin PFZ^k}{\operatorname{argmin}} \left(dist(\mathbf{v'}_{rob} \cdot t_h, \; \mathbf{g}_{rob}) \right). \tag{16}$$

This equation implies that our navigation method deviates the robot by the least amount from its goal such that it does not enter any group's PFZ. However, in dense scenarios, when the robot encounters many potentially freezing groups and their corresponding PFZs, Equation 16 may not be able to compute a feasible solution for ϕ_{com}. In such cases, a potential solution is to let the robot pass through the PFZ of a low cohesion group (see Fig. 5a). In such cases, we formulate the deviation angle as,

$$\phi_{com} = \underset{\mathbf{v'}_{rob} \cdot t_h \in \mathcal{P}}{\operatorname{argmin}} \left(dist(\mathbf{v'}_{rob} \cdot t_h, \; \mathbf{g}_{rob}) \right), \tag{17}$$

$$\mathcal{P} = PFZ^{min} \setminus (PFZ^{min} \cap PFZ^k) \; \forall k \in \{1, 2, ..., M.\}, \tag{18}$$

where PFZ^{min} is the PFZ of the group with the minimum cohesion in the scene. Since the permeability of low cohesion groups is high, the above formulation also lowers the probability of freezing.

Proposition V.1. *The deviation angles computed by CoMet-based navigation (Equations 16 or 18), and Frozone (2) satisfy the relationship* $\phi_{froz} \geq \phi_{com}$.

Proof. Consider the scenario shown in Fig. 3, with PFZ_{froz} depicted as a 5-sided convex polygon. A vertical line segment connects the robot (O_{rob}) to its goal and the robot's deviation is measured relative to it. Frozone deviates $\mathbf{v}_{rob}$ such that this point lies on the boundary of PFZ_{froz} (Equation 3) or in the exterior of PFZ_{froz} (Equation 4), depending on whichever leads to a lower deviation. Let $O_{froz} = [x_{froz}, y_{froz}]$ be the point on the boundary of PFZ_{froz} to which Frozone deviates (by angle ϕ_{froz}) the robot. Based on Equations 16 and 18, our CoMet-based solution deviates the robot to a point in $PFZ_{froz} \setminus \bigcup_{k=1}^{M} PFZ^k$.

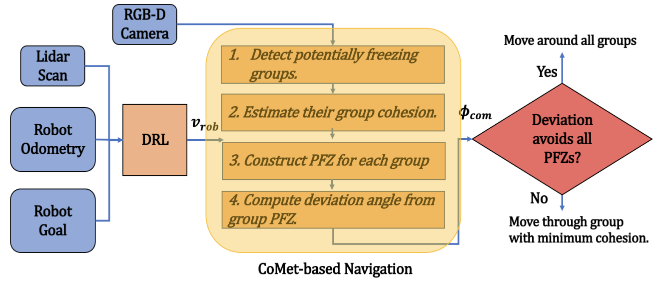

Fig. 4: Our Socially-Compliant Navigation Algorithm: We use a DRL framework used to guide the robot to its goal and handle static obstacles. CoMet-based navigation considers each group in the scene, identifies groups which could result in FRP, constructs group Potential Freezing Zones (PFZ), and computes a deviation angle to avoid such zones. Our formulation results in lower occurrence of freezing, lower deviations for the robot with respect to pedestrians and groups, and avoiding high cohesion groups by moving around them. In dense scenarios, when there is no feasible solution for the deviation angle, our method navigates the robot through the group with the lowest cohesion. All these behaviors improve the naturalness of the robot trajectory's.

Since, $\bigcup_{k=1}^{M} PFZ^k \subseteq PFZ_{froz}$, there exists a set $\mathbf{F} \subseteq PFZ_{froz} \setminus \bigcup_{k=1}^{M} PFZ^k$ such that all positions in $\mathbf{F}$ lead to a deviation angle $\phi \leq \phi_{froz}$. For instance, in Fig. 3, $\mathbf{F}$ can be a set just within the boundary of PFZ_{froz} with Y-coordinates greater than y_{froz}. Since Equations (16) and (18) optimize for minimum deviation from the goal, $\phi_{com} \in \mathbf{F}$. This implies that $\phi_{com} \leq \phi_{froz}$. The equality holds when $\mathbf{F} = \emptyset$ or when the closest edge of PFZ_{froz} to the robot corresponds to the PFZ of a group. Based on the triangle inequality, shorter deviations lead to shorter trajectory lengths. ∎

This bound also guarantees that our new navigation algorithm generates trajectory that are more natural, as compared to Frozone [6]. We integrate our CoMet-based navigation method with a DRL-based navigation scheme[28]. Figure 4) shows the components of our navigation algorithm that is used to compute trajectories that are more natural.

VI. Implementation and Results

In this section, we describe our implementation of computing group cohesion and socially-compliant navigation. We then evaluate CoMet in different standard pedestrian datasets that are annotated with perceived group cohesion levels. We highlight the performance of our navigation algorithm and show benefits over prior methods in terms of the following metrics: freezing rate, deviation angle, and normalized path-length.

A. Implementation

In order to evaluate CoMet, we annotate groups in pedestrian datasets such as MOT, KITTI, ETH etc, as low, medium and high cohesion groups using multiple human annotators. These annotated datasets are used as the ground truth. We choose these datasets since they depict groups in real-world scenarios with various lighting conditions, crowd density and occlusions. We manually tune the weighing constants in the CoMet formulation (K_p, K_w, K_s, K_i) and set thresholds on the cohesion score to classify groups into the aforementioned categories based on the annotations in one of the datasets.

(a) Groups standing in a corridor (b) Group standing in a line (c) Group sitting in a common area

Fig. 5: Qualitative evaluations of the trajectories generated using our algorithm (shown as green) in different scenarios. We also compare with the trajectories generated using Frozone [6](shown as yellow) and a DRL-based algorithm [28](shown as orange). Each group's PFZ is shown as a red region on the floor. We evaluate our method in three different real-world scenarios with tight spaces, with people standing or sitting. Our method differentiates between low (in red) and high cohesion (in green) groups, and navigates only between low cohesion groups. Frozone algorithm behaves in a conservative manner and halts the robot in dense scenarios. DRL [28] prioritizes moving towards the goal and passes through high cohesion groups (see (a)). Overall, our approach results in socially-compliant trajectories.

We evaluate CoMet's precision and recall in these groups in all other datasets. We have evaluated our algorithm on a Turtlebot 2 robot, mounted with an Intel RealSense RGB-D camera (for pedestrian tracking and localization), and a 2-D Hokuyo lidar.

B. Analysis

CoMet Classification: Table II highlights CoMet's classification precision and recall in multiple datasets. CoMet observes pedestrians in these datasets for ~ 5 seconds. During this period, it is able to update its initial classification as pedestrian's trajectories change. For instance, a group initially perceived as high cohesion may have its members move apart and is thereby classified as a low cohesion group. Moreover, it is easier to detect features corresponding to proximity and walking speed, as compared to interaction between the pedestrians. An interesting observation is that human annotators tend to classify groups in extremes, i.e., as either high cohesion or low cohesion groups. This leads to low number of data points for medium-cohesion groups. This ground truth observation affects the effectiveness of our approach.

Socially-Compliant Navigation: We compute our methods's trajectories versus Frozone's and a DRL algorithm [28] trajectories qualitatively. We highlight the differences in Fig. 5. Our approach is able to identify low cohesion groups successfully and navigate through them without interfering with high cohesion groups. In contrast, Frozone halts the robot completely, since it does not assume pedestrian cooperation in its formulation. The DRL method prioritizes reaching the goal with the minimum path length and therefore navigates through groups regardless of their cohesion. Therefore, the DRL algorithm can generate obtrusive trajectories.

We also compare our navigation algorithm with Frozone in simulated environments with varying number $(10-50)$ of pedestrians in a corridor-like scenario. Pedestrians are given random initial locations, based on which they are classified into groups and PFZ_{froz} and group PFZs (PFZ^k) are computed. The robot needs to navigate through the pedestrians to reach its goal. We use the following metrics: (1) *Average deviation angle* measured relative to the line connecting the start and goal locations. (2) *Freezing Rate* measured as the number of times the robot halted/froze over the total number of trials, and (3) *Normalized Path Length* measured as the robot's path length over the length of the line connecting the start and goal locations.

Dataset Video	Low-Cohesion	Medium-Cohesion	High-Cohesion
ADL-Rundle	1.00/1.00	-	1.00/0.75
AVG-TownCenter	1.00/1.00	0.50/0.67	0.50/0.33
ETH Jelmoli	1.00/1.00	-	1.00/0.80
ETH Bahnhof	1.00/1.00	1.00/0.67	0.80/1.00
KITTI-16	1.00/1.00	1.00/0.50	0.33/1.00
KITTI-17	1.00/1.00	-	1.00/0.667
MOT17-08	1.00/1.00	-	1.00/0.50
MOT17-11	1.00/1.00	0.75/1.00	1.00/0.833
MOT20-02	1.00/1.00	0.34/0.33	0.50/0.33
TUD	1.00/0.857	-	1.00/1.00
Venice	0.875/0.875	-	0.938/0.60

TABLE II: Table shows the precision (first value in each column) and recall (second) values for the three classes when CoMet observes groups in the video for ~ 5 seconds. CoMet's parameters have been tuned based on the ground truth in one of the datasets. We observe good accuracy for high-cohesion and low-cohesion groups. The original datasets have fewer occurrences of medium-cohesion groups, which impacts our precision.

Metrics	Method	10 peds	20 peds	30 peds	40 peds	50 peds
Avg. Deviation Angle (lower better)	Frozone + DRL Our Method + DRL	46.41 **41.07**	46.67 **44.31**	42.85 **41.11**	50.80 **37.47**	53.22 **34.19**
Freezing Rate (lower better)	Frozone + DRL Our Method + DRL	0 0	0.29 0	0.25 0	0.43 0	0.57 0
Normalized Path Length (lower better)	Frozone + DRL Our Method + DRL	1.46 **1.12**	1.45 **1.29**	1.55 **1.27**	1.59 **1.35**	1.51 **1.35**

TABLE III: We compare different navigation metrics computed using our CoMet-based navigation algorithm relative to Frozone [6]. We observe that our method consistently results in lower values corresponding to all these metrics. This signifies improved naturalness of the robot's trajectory computed using our approach.

Our method results in lower values with respect to all these metrics, as compared to Frozone on the same scenarios. As the crowd size, density or the number of groups increase, Frozone's conservative formulation makes the robot freeze at a much high rate. From this analysis, we observe that our method produces trajectories with high social compliance and naturalness. Furthermore, our approach significantly reduce the occurence of freezing behavior.

VII. Conclusions, Limitations and Future Work

We present a novel method to compute the cohesion of a group of people in a crowd using visual features. We use our cohesion metric to design a novel robot navigation algorithm that results in socially-compliant trajectories. We highlight the benefits over previous algorithms in terms of following metrics: reduced freezing, deviation angles, and path lengths. We test our cohesion metric in annotated datasets and observe a high precision and recall.

Our method has some limitations. We model cohesion through a linear relationship between the features, which may not work in all scenarios. In addition, there are other characteristics used to estimate cohesion, including age, gender, environmental context, cultural factors, etc. that we do not take into account. Our approach also depends on the accuracy of how these features are detected, which may be affected due to lighting conditions and occlusions. Our navigation assumes that different groups in a group exhibit varying cohesion, which may not hold all the time. As part of future work, we hope to address these limitations and evaluate our approach in crowded real-world scenes.

REFERENCES

[1] P. Trautman and A. Krause, "Unfreezing the robot: Navigation in dense, interacting crowds," in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Oct 2010, pp. 797–803.

[2] L. He, J. Pan, S. Narang, and D. Manocha, "Dynamic group behaviors for interactive crowd simulation," in *Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation*. Eurographics Association, 2016, p. 139–147.

[3] S. Bandini, A. Gorrini, L. Manenti, and G. Vizzari, "Crowd and pedestrian dynamics: Empirical investigation and simulation," in *Proceedings of Measuring Behavior*, 08 2012.

[4] Y. F. Chen, M. Everett, M. Liu, and J. How, "Socially aware motion planning with deep reinforcement learning," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 09 2017, pp. 1343–1350.

[5] L. Tai, J. Zhang, M. Liu, and W. Burgard, "Socially compliant navigation through raw depth inputs with generative adversarial imitation learning," in *ICRA*, May 2018, pp. 1111–1117.

[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] Y. F. Chen, M. Liu, M. Everett, and J. P. How, "Decentralized non-communicating multiagent collision avoidance with deep reinforcement learning," in *ICRA*. IEEE, 2017, pp. 285–292.

[8] A. Jagan Sathyamoorthy, J. Liang, U. Patel, T. Guan, R. Chandra, and D. Manocha, "DenseCAvoid: Real-time Navigation in Dense Crowds using Anticipatory Behaviors," *arXiv e-prints*, p. arXiv:2002.03038, Feb. 2020.

[9] N. Ashton, M. E. Shaw, and A. P. Worsham, "Affective reactions to interpersonal distances by friends and strangers," *Bulletin of the psychonomic society*, vol. 15, pp. 306–308, 1980.

[10] S. Karau and K. Williams, "The effects of group cohesiveness on social loafing and social compensation." *Group Dynamics: Theory, Research, and Practice*, vol. 1, pp. 156–168, 1997.

[11] E. Knowles, "Boundaries around group interaction: the effect of group size and member status on boundary permeability," *Journal of personality and social psychology*, vol. 26, no. 3, p. 327—331, June 1973. [Online]. Available: https://doi.org/10.1037/h0034464

[12] E. T. Hall, R. L. Birdwhistell, B. Bock, P. Bohannan, A. R. Diebold, M. Durbin, M. S. Edmonson, J. L. Fischer, D. Hymes, S. T. Kimball, W. La Barre, , J. E. McClellan, D. S. Marshall, G. B. Milner, H. B. Sarles, G. L. Trager, and A. P. Vayda, "Proxemics [and comments and replies]," *Current Anthropology*, vol. 9, no. 2/3, pp. 83–108, 1968. [Online]. Available: https://doi.org/10.1086/200975

[13] J. Wagnild and C. M. Wall-Scheffler, "Energetic consequences of human sociality: Walking speed choices among friendly dyads," *PLOS ONE*, vol. 8, no. 10, pp. 1–6, 10 2013. [Online]. Available: https://doi.org/10.1371/journal.pone.0076576

[14] J. Shao, C. C. Loy, and X. Wang, "Scene-independent group profiling in crowd," in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 2227–2234.

[15] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *CoRR*, vol. abs/1804.02767, 2018. [Online]. Available: http://arxiv.org/abs/1804.02767

[16] R. Chandra, U. Bhattacharya, A. Bera, and D. Manocha, "Densepeds: Pedestrian tracking in dense crowds using front-rvo and sparse features," *arXiv preprint arXiv:1906.10313*, 2019.

[17] W. Choi, Y.-W. Chao, C. Pantofaru, and S. Savarese, "Discovering groups of people in images," in *Computer Vision – ECCV 2014*, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds. Cham: Springer International Publishing, 2014, pp. 417–433.

[18] A. K. Chandran, L. A. Poh, and P. Vadakkepat, "Identifying social groups in pedestrian crowd videos," in *2015 Eighth International Conference on Advances in Pattern Recognition (ICAPR)*, 2015, pp. 1–6.

[19] R. M. Favaretto, L. L. Dihl, and S. R. Musse, "Detecting crowd features in video sequences," in *2016 29th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, 2016, pp. 201–208.

[20] Z. Ebrahimpour, W. Wan, O. Cervantes, T. Luo, and H. Ullah, "Comparison of main approaches for extracting behavior features from crowd flow analysis," *ISPRS International Journal of Geo-Information*, vol. 8, no. 10, 2019. [Online]. Available: https://www.mdpi.com/2220-9964/8/10/440

[21] M. Yang, L. Rashidi, A. S. Rao, S. Rajasegarar, M. Ganji, M. Palaniswami, and C. Leckie, "Cluster-based crowd movement behavior detection," in *2018 Digital Image Computing: Techniques and Applications (DICTA)*, 2018, pp. 1–8.

[22] M. Cristani, L. Bazzani, G. Paggetti, A. Fossati, D. Tosato, A. Del Bue, G. Menegaz, and V. Murino, "Social interaction discovery by statistical analysis of f-formations," in *Proceedings of the British Machine Vision Conference*. BMVA Press, 2011, pp. 23.1–23.12, http://dx.doi.org/10.5244/C.25.23.

[23] T. Kruse, A. K. Pandey, R. Alami, and A. Kirsch, "Human-aware robot navigation: A survey," *Robotics Auton. Syst.*, vol. 61, pp. 1726–1743, 2013.

[24] C. Chen, Y. Liu, S. Kreiss, and A. Alahi, "Crowd-robot interaction: Crowd-aware robot navigation with attention-based deep reinforcement learning," in *ICRA*. IEEE, 2019, pp. 6015–6022.

[25] A. Bera, T. Randhavane, R. Prinja, and D. Manocha, "Sociosense: Robot navigation amongst pedestrians with social and psychological constraints," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Sep. 2017, pp. 7018–7025.

[26] L. Liu, D. Dugas, G. Cesari, R. Siegwart, and R. Dubé, "Robot navigation in crowded environments using deep reinforcement learning," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020, pp. 5671–5677.

[27] M. Everett, Y. F. Chen, and J. P. How, "Motion planning among dynamic, decision-making agents with deep reinforcement learning," in *IROS*. IEEE, 2018, pp. 3052–3059.

[28] P. Long, T. Fan, X. Liao, W. Liu, H. Zhang, and J. Pan, "Towards Optimally Decentralized Multi-Robot Collision Avoidance via Deep Reinforcement Learning," *arXiv e-prints*, p. arXiv:1709.10082, Sep 2017.

[29] K. M. Kitani, B. D. Ziebart, J. A. Bagnell, and M. Hebert, "Activity forecasting," in *Computer Vision – ECCV 2012*, A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, and C. Schmid, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 201–214.

[30] M. Pfeiffer, U. Schwesinger, H. Sommer, E. Galceran, and R. Siegwart, "Predicting actions to act predictably: Cooperative partial motion planning with maximum entropy models," 10 2016, pp. 2096–2101.

[31] H. Kretzschmar, M. Spies, C. Sprunk, and W. Burgard, "Socially compliant mobile robot navigation via inverse reinforcement learning," *The International Journal of Robotics Research*, vol. 35, no. 11, pp. 1289–1307, 2016. [Online]. Available: https://doi.org/10.1177/0278364915619772

[32] T. Randhavane, A. Bera, E. Kubin, A. Wang, K. Gray, and D. Manocha, "Pedestrian dominance modeling for socially-aware robot navigation," in *2019 International Conference on Robotics and Automation (ICRA)*, 2018.

[33] A. Bera, T. Randhavane, A. Wang, D. Manocha, E. Kubin, and K. Gray, "Classifying group emotions for socially-aware autonomous vehicle navigation," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2018.

[34] E. Goffman, *Encounters: Two studies in the sociology of interaction*. Ravenio Books, 1961.

[35] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," *arXiv e-prints*, p. arXiv:1506.02640, June 2015.

[36] N. Wojke, A. Bewley, and D. Paulus, "Simple Online and Realtime Tracking with a Deep Association Metric," *arXiv e-prints*, p. arXiv:1703.07402, Mar. 2017.

[37] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME–Journal of Basic Engineering*, vol. 82, no. Series D, pp. 35–45, 1960.

[38] T. Baltrusaitis, A. Zadeh, Y. C. Lim, and L. Morency, "Openface 2.0: Facial behavior analysis toolkit," in *2018 13th IEEE International Conference on Automatic Face Gesture Recognition (FG 2018)*, 2018, pp. 59–66.

[39] J. Van Den Berg, S. Guy, M. Lin, and D. Manocha, "Reciprocal n-body collision avoidance," in *Robotics Research - The 14th International Symposium ISRR*, ser. Springer Tracts in Advanced Robotics, no. STAR, June 2011, pp. 3–19, 14th International Symposium of Robotic Research, ISRR 2009 ; Conference date: 31-08-2009 Through 03-09-2009.