Modeling and Identification of Group Motion via Compound Evaluation of Positional and Directional Cues

Zeynep Yücel  Takahiro Miyashita  Norihiro Hagita
Intelligent Robotics and Communication Laboratories, ATR, Japan
zeynep@atr.jp, miyasita@atr.jp, hagita@atr.jp

Abstract

This paper addresses the problem of identification of pedestrian groups in crowded environments. To that end, positional and directional relations are modeled accounting for different environmental features and group configurations. Subsequently, a pair of simultaneously observed pedestrians is identified to belong to the same group or not utilizing these models in a parallel manner, which defines a compound hypothesis testing scheme. In case of ambiguities, local and global indicators of group relation are employed in quantifying the reliabilities of the two individual decisions. The contribution of this study lies in the improvement in positional and directional relation models to adjust to different environments and group configurations, description of compound evaluation, as well as resolution of ambiguities proposing uncertainty measures based on local and global indicators of group relation.

1 Introduction and motivation

Public environments are often monitored bearing an immense amount of data. In order to attain a clear understanding of this information, it is desirable to utilize automatic interpretation methods, which are resilient against the inherent contextual and practical asperities of the data. In that respect, treating the observed crowd behavior as a collection of interactions between individual pedestrians is suggested to lead to a significant loss of information [1]. This is due to the fact that people who are engaged in a social relation present local cohesive behavior, which makes the independent individual behavior assumption ineffective. Understanding such group relation offers potential improvement in effective modeling of human locomotion in crowded settings and interaction of groups [2], detection of collective unusual behavior [3], and realistic crowd simulations as well as substantial data association performance and enhancement of tracking [4].

The outline of the paper is as follows. Section 2 elaborates on the properties of the datasets employed in modeling and introduces the notation and the overview of the proposed method. Sections 3 and 4 define the motion models. Section 5 explains the integration of individual indicators and defines uncertainty measures. Finally, Section 6 presents the experimental results addressing performance, generalization and sensitivity and provides a comparison to an earlier work.

2 Datasets and overview of method

Various environmental conditions and group structures are considered in identification of group relation. In order to demonstrate the flexibility of the method against changes in these conditions, the BIWI and APT datasets, which assume different environmental attributes and distinctions in group structure, are employed [5, 6]. BIWI Walking Pedestrians dataset is composed of two experiments recorded from birds-eye view with a total of 650 tracks over 25 minutes [5]. The first experiment, BIWI-ETH, is performed in the entrance passage of a building with a strong flow of pedestrians between the street and the building (see Figure 1-(a)). The second experiment, BIWI-Hotel, is recorded over a sidewalk, where there is a remarkable flux along (see Figure 1-(b)). In this study, such scenes with an uneven distribution of pedestrian flow are called non-uniform environments. APT Pedestrian Behavior Analysis dataset is recorded in the entrance hall of a shopping center [7] (see Figure 1-(c)). Unlike BIWI, there is no prominent flow in any direction. Due to the homogeneous distribution of flow, APT dataset is regarded to come from a uniform environment. In addition to the distinction between existence of preferred direction, the
two sets present significant difference in group structure as well. BIWI involves various multi-partner groups (3 to 6 people), whereas APT contains people moving individually or in groups of 2.

In this framework, two pedestrians, who are observed simultaneously, are termed as a pair. Following the definition of group given in [8], the pairs of pedestrians who belong to the same group such as \( \{p_i, p_j\} \) of Figure 1-(a), constitute the set \( \mathcal{G} \), whereas the pairs of pedestrians who do not belong to the same group such as \( \{p_i, p_h\} \) or \( \{p_i, p_m\} \) comprise the complementary set \( \mathcal{G}^c \). Within \( \mathcal{G} \), the degree of neighborhood is defined by the configuration of the group members. In that respect, the people, who do not have any other pedestrian significantly close to the axis connecting them, are regarded as first neighbors. For instance, \( \{p_i, p_j\} \) are first neighbors, whereas \( \{p_i, p_k\} \) are second neighbors.

Once a pair is observed, the question of interest is whether it belongs to \( \mathcal{G} \) or \( \mathcal{G}^c \). To that end, two sorts of observations are utilized, namely the positional and directional relations. The positional relation \( \Delta \) is expressed by the set of absolute values of the displacements \( \{\delta\} \), whereas the directional relation \( \Theta \) is expressed by the set of angles between velocity vectors \( \{\theta\} \) (see Figure 1-(a)). Both sets involve measurements collected throughout the duration of the observation of the pair. Moreover, each of \( \mathcal{G} \) and \( \mathcal{G}^c \) is described by two models characterizing the positional and directional relations, i.e. \((\Delta_{\mathcal{G}}, \Theta_{\mathcal{G}})\) or \((\Delta_{\mathcal{G}^c}, \Theta_{\mathcal{G}^c})\). The problem is deliberated with two different applications of the same approach in parallel, i.e. investigating whether \( \Delta \sim \Delta_{\mathcal{G}} \) or \( \Delta \sim \Delta_{\mathcal{G}^c} \) and \( \Theta \sim \Theta_{\mathcal{G}} \) or \( \Theta \sim \Theta_{\mathcal{G}^c} \). The final decision is rendered based on the outcomes of these two, where the outcome implicating a lower uncertainty is preferred in case of ambiguities.

3 Motion model of \( \mathcal{G} \)

The positional relation of a pair \( \{p_m, p_n\} \in \mathcal{G} \), where \( p_m \) and \( p_n \) are first neighbors as in Figure 1-(a), is expected to be described by a distribution over a short range with a single peak centered around a comfortable distance between \( p_m \) and \( p_n \). This suggests a Rice distribution,

\[
\Delta_{\mathcal{G}}(\delta|\mu, \sigma) = \frac{\delta}{\sigma^2} \exp \left( -\frac{\delta^2 - \mu^2}{2\sigma^2} \right) I_0 \left( \frac{\delta \mu}{\sigma^2} \right),
\]

where \( \nu \) represents the comfortable distance and \( I_0 \) is the modified Bessel function of first kind with order 0 [9].

Moreover, in multi-partner groups the distance between the pedestrians who are not first neighbors such as \( p_i \) and \( p_k \) in Figure 1-(a), is considered to be described by a convolution of Rice distributions. Namely, for a multi-partner group composed of \( N + 1 \) people, this multimodal framework implies \( \Delta_{\mathcal{G}} = \sum_{i=1}^{N} K_i \Delta_{\mathcal{G}_i} \), where \( K_i \) is the observation frequency of \( i^{th} \) neighborhood. \( \Delta_{\mathcal{G}_i} \) denotes the distribution between \( i^{th} \) neighbors and is equivalent to convolution of \( \Delta_{\mathcal{G}} \) with itself \( i \) times. It is suggested to restrict \( N \in \{1, 2, 3\} \) to conserve normal distribution assumption. Moreover, small groups often tend to keep an abreast formation in relatively low pedestrian densities, whereas larger groups of 4 or more people can be arranged in more complex configurations [10], which eliminates the need to extend \( N \) in the first place.

On the other hand, the directional relation of a pair of pedestrians \( \{p_m, p_n\} \in \mathcal{G} \) is expected to be described by aligned velocity vectors, i.e. \( \mu(\theta) = 0 \) as \( \mu \) denotes mean value. Moreover, this alignment is suggested to be independent of their degree of neighborhood or whether the environment is uniform or nonuniform. This is due to the fact that the directional correlation regarding \( \mathcal{G} \) depends mostly on the short range events. Thus, due to its periodic properties \( \theta \) is modeled as a von Mises distribution.
distribution [11], which is the circular analogue of the Gaussian distribution,
\[ \Theta_G(\theta|\kappa) = \frac{\exp(\kappa \cos(\theta))}{2\pi I_0(\kappa)}, \] (2)
where \( \kappa \) is analogous to \( 1/\sigma^2 \) of normal distribution.

4 Motion model of \( \bar{G} \)

The positional relation of \( \{p_i, p_h\} \in \bar{G} \) is modeled based on the assumption that the relative locations are independent due to the lack of social relation. This assumption makes the problem equivalent to selecting two points randomly in the observation environment and measuring the distance in between. Suppose that the dimensions of the observation environment along \( x- \) and \( y- \) axes are \( D \). Then,

\[ \Delta_G(\delta) = \begin{cases} \frac{1}{2} 2\delta (\delta^2 - 4\delta + \pi), & 0 < \delta < D, \\ \frac{1}{2} 2\delta \sqrt{\delta^2 - 1} - (\delta^2 + 2 - \pi) - 4 \tan^{-1} (\sqrt{\delta^2 - 1}), & D < \delta < D\sqrt{2}. \end{cases} \] (3)

The directional relation of \( \{p_i, p_h\} \in \bar{G} \) is expected to be uncorrelated provided that they are free to move in any direction, i.e. in uniform environments. Thus we can assume \( \mu(\theta) = 0 \). Hence, \( \Theta_{\bar{G}} \) regarding uniform environments is modeled as in Equation 2, where the parameter \( \kappa \) enables modeling of different behavior.

In nonuniform environments, the flow leads to a directional correlation even though the pedestrians do not belong to the same group. In our data, \( \mu(\theta) \) concerning \( \bar{G} \) is either 0 or \( \pi \), since there is a single dominant flux as in Figure 1-(b). Thus, assuming that the flow is as strong in one direction as in the other, the distribution of \( \theta \) regarding \( \bar{G} \) can be approximated with an equally weighted linear combination of two von Mises distributions with \( \mu(\theta) = 0 \) and \( \mu(\theta) = \pi \),

\[ \Theta_{\bar{G}}(\theta|\kappa) = \frac{\exp(\kappa \cos(\theta))}{4\pi I_0(\kappa)} + \frac{\exp(\kappa \cos(\theta - \pi))}{4\pi I_0(\kappa)}. \] (4)

5 Hypothesis testing

Since the individual observations, \( \Delta \) and \( \Theta \), satisfy the mutual exclusiveness and complementarity conditions, the problem is treated in a coupled hypothesis testing scheme. In this manner, as long as the outcomes of the tests agree, a decision can be made confidently. In case of conflicting decisions, a measure of uncertainty needs to be defined to resolve for the final decision. In what follows we describe the way in which the individual decisions are carried out and we define the measures of uncertainty employed in resolving the final decision in case of contradictions.

In binary decisions, likelihood ratio test is one way of determining the underlying model. Consider the decision of whether \( \Delta \) comes from \( \Delta_G \) or \( \Delta_{\bar{G}} \). The log-likelihood ratio test suggests that \( \Delta \sim \Delta_{\bar{G}} \) if \( L^\delta > 0 \), where \( L^\delta \) denotes the log-likelihood ratio of being in a group relation over not being in a group relation. Otherwise, \( \Delta \sim \Delta_G \). The decision based on \( \theta \) is carried out in a similar manner.

As long as \( L^\delta \) and \( L^\theta \) have the same sign, a confident decision is made regarding the group relation of the pair of interest. However, contradictions might arise in cases like crossing next to each other, moving along a flow, or passing through passages. One straightforward resolution to this contradiction is to select the decision which has a larger absolute value of log-likelihood ratio. We claim this naive approach is not efficient in resolving the contradiction. Instead we propose the following reliability based approach.

In order to compensate for the effect of these misleading cues, an uncertainty measure is devised. Inspired from the Kullback-Leibler divergence, a reliability estimate is employed in quantifying the uncertainty of individual decisions. The Kullback-Leibler divergence of two distributions such as \( P \) and \( Q \), is defined as,

\[ D_{KL}(P||Q) = \sum_i p(i) \log \left( \frac{p(i)}{q(i)} \right). \] (5)

Note that this measure is not symmetric, i.e. \( D_{KL}(Q||P) \neq D_{KL}(P||Q) \). Hence, mathematically speaking it is not a distance measure but it quantifies the difference between two distributions. In order to have a common reference point, the divergence terms are computed with respect to the observed distributions. Hence, the divergences relating \( \delta \) with respect to \( \bar{G} \) and \( \bar{G} \) are defined as \( D_G^\delta = D_{KL}(\Delta||\Delta_G) \) and \( D_{\bar{G}}^\delta = D_{KL}(\Delta||\Delta_{\bar{G}}) \). Since they embrace all \{\delta\} through the summation term, we call them global indicators of group motion.

However, \( \theta \) relating \( G \) does not present a behavior as regular as \( \delta \) of \( \bar{G} \). Thus, it is proposed to focus on its local characteristics so as to avoid the misleading temporal imperfections that might lead to a false similarity to \( \bar{G} \). Namely, the divergence term relating \( \theta \) with respect to \( \bar{G} \) is defined as,

\[ D_{\bar{G}}^\theta(\theta||\Theta_{\bar{G}}) = \max_{\theta} \left\{ \Theta(\theta) \log \left( \frac{\Theta(\theta)}{\Theta_{\bar{G}}(\theta|\kappa)} \right) \right\}, \] (6)

where the divergence of \( \theta \) with respect to \( \bar{G} \) is computed in a similar manner. This implies that only the diver-
gence value which indicates the maximum dissimilarity is addressed defining a local indicator of group motion.

A direct comparison of these four divergence terms is not possible since they are not defined in the same reference frame. Hence, two uncertainty measures are defined regarding each individual decision as the ratio of the concerning divergence values,

\[ \rho^\delta = \frac{D^\delta_G}{D^\delta_{\bar{G}}}, \]
\[ \rho^\theta = \frac{D^\theta_G}{D^\theta_{\bar{G}}}. \] (7)

The final resolution is determined by picking the decision with lower uncertainty.

6 Experimental results and discussion

The model parameters are solved by randomly selecting 10% of all the concerning samples and minimizing the squared error between this distribution and the proposed models. For brevity’s sake, examples of only a few cases are illustrated in Figure 2. Figure 2-(a) ascertains that the multimodal framework captures the peaks concerning each neighborhood. Moreover, \( \delta \) of \( \bar{G} \) is modeled by the non-parametric approach of Equation 3 with a satisfactorily good performance as in Figure 2-(b). The model of \( \theta \) regarding \( G \) has a sharp peak around 0 in Figure 2-(c) as expected. On the other hand, the model concerning \( \bar{G} \) presents a clear concentration around 0 and \( \pi \) due to flow characteristics.

The shuffling and random selection procedure is repeated 100 times and the variation of model parameters is quite insignificant indicating the resiliency against varying training sets.

At each of these 100 runs, the resolved model parameters coming from the training set constituting 10% of the data, are employed in identification of the group relation concerning the remaining 90%. The detection rates are given in Table 1.

In order to demonstrate the improvement introduced by the hypothesis testing scheme explained in Section 5 to the motion models described in Sections 3 and 4, we pick the decision with larger absolute value of log-likelihood ratio in case of contradictory decisions. The performance of this approach is given in the first column of Table 1 under maximum absolute log-likelihood ratio. Comparing the first and second columns of Table 1, it is concluded that particularly in non-uniform environments such an approach is quite naive and is not able to distinguish the correlation due to the flow from the correlation due to group relation.

Besides, our method outperforms [6] in terms of overall detection rates of both datasets. Moreover, our identification rates of \( G \) and \( \bar{G} \) do not present any significant distinction for either dataset. This indicates that our method does not present any bias in favor of a particular class, unlike [6] and the method of maximum absolute log-likelihood. Both of these methods detect \( G \) with a poor detection rate for BIWI dataset implying a positive bias for \( \bar{G} \) and increasing the overall detection rate misleadingly. This shortcoming of [6] and method of maximum absolute log-likelihood is due to the fact that the nonuniform environment and high pedestrian density of BIWI are not accounted for in these methods. However, we adjust our models according to the environmental conditions and groups structure, which enables competent identification rates in various scenarios.

Figure 2. The observed and modeled distributions of \( \delta \) regarding (a) \( G \) of BIWI and (b) \( \bar{G} \) of APT, and \( \theta \) regarding (c) \( G \) and (d) \( \bar{G} \) of BIWI.
Table 1. Comparison of performance.

<table>
<thead>
<tr>
<th>Method of <a href="%25">6</a></th>
<th>Maximum absolute log-likelihood ratio(%)</th>
<th>Proposed method(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>BIWI</td>
<td>71.69</td>
<td>97.06</td>
</tr>
<tr>
<td>APT</td>
<td>89.81</td>
<td>88.08</td>
</tr>
</tbody>
</table>

7 Conclusion

In this study, positional and directional models are proposed for identification of pedestrian groups in crowded environments together with a compound evaluation scheme. Different environmental characteristics are accounted for in addition to varying group structures. Our results indicate that the proposed models grasp the characterizing features of different environmental settings and varying patterns of group relation. Moreover, the model parameters are shown to be stably derived from a small set of data. Besides, the group relations are illustrated to be identified with satisfactorily high rates. The efficacy of compound evaluation is verified by a comparison to another method in literature. Finally, our contributions are listed as the improvement in positional and directional models to adjust to different environments and group structures, description of compound evaluation, and resolution of ambiguities by the proposed uncertainty measures based on local and global indicators.

References