Modeling Indicators of Coherent Motion

Zeynep Yücel   Francesco Zanlungo   Tetsushi Ikeda   Takahiro Miyashita   Norihiro Hagita

Abstract— This study focuses on joint motion patterns of humans that move together with other humans or objects. Since this scope embraces 'group motion', which relates only humans, and expands its extent of interactions accounting for various auxiliary instruments such as walking aids or pushcarts, we term this collective motion pattern as 'coherent' motion. Coherence is proposed to be characterized by the distance between the moving parties, the scalar product of their velocities and the scalar product of the velocity vector and the displacement vector. The contribution of this study lies in the formulation of coherence in terms of the listed features through explicit mathematical models. The models are developed in accordance with a large database recorded in an uncontrolled environment involving a total of more than 500 mobile entities. The performance of the proposed models is evaluated qualitatively by comparing them to the empirical data and quantitatively by employing log-likelihoods. Comparison to an earlier work indicates that the proposed models improve the identification of coherence quality significantly well.

I. Introduction and Motivation

Current day robotics has evolved from early industrial robots designed for automatically performing specific tasks towards robots with extended autonomous capabilities and decision making skills as well as a potential for adaptation against varying environmental conditions and social interaction. In that respect, the development of flexible architectures, which are capable of interpreting the sensory information in an intelligent manner to account for semantic relations, emerges as a crucial matter. Nevertheless, most robotic agents address the issue of adaptation against varying requirements from a goal-oriented point of view, which leads to an abstraction of the constituent elements of the environment from the mutual relation. This implies that understanding the interaction of these elements with the environment and with each other still stands as an open field.

However, an intelligent conception of all the environmental elements requires exact attribution of situational, social, and functional qualities. This defines a very complex task bearing a profound training stage integrated with a wide range of sensory information. So as to accommodate this goal, investigation of the human element of the environment adopting an interaction-orientated standpoint is proposed to be significantly availing from a practical perspective. The potential fields, where attribution of interrelated qualities to humans offer a facilitating means, involve human-robot interaction, multi-object tracking, activity recognition, and

statistical analysis of pedestrian behavior. The interactions which present a significant importance for those applications span a wide range of social and functional relations. In that respect, this study focuses in particular on certain motion characteristics of humans and several mobile objects and makes inferences about the joint motion patterns.
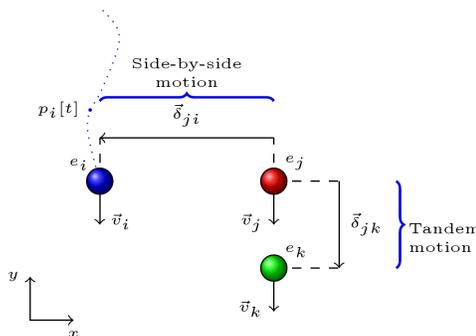

Fig. 1. Examples for coherent pairs. Three entities are depicted with ideal coherence characteristics of perfectly aligned velocity vectors.

Specifically the entities presenting considerable correlation in their trajectory and speed are identified. We call any two entities, which are observed simultaneously, a *pair*. The pairs which move along a similar trajectory with a short distance and similar velocities are termed as *coherent* pairs. On the contrary, the entities, which do not have a correlation in trajectory or velocity, and happen to be observed simultaneously, are named *non-coherent* pairs. Among the pairs of interest, the possible configurations of coherent motion are *side-by-side* and *tandem* (see Figure 1). Side-by-side motion is defined as the motion of two entities which walk abreast with a common goal and social connection, whereas tandem motion is defined as the motion of a pair consisting of two entities arranged one behind the other.

As the motion characteristics of listed types of entities are observed, the behavioral distinctions are concluded to be most prominent in the distance and relative velocity patterns. In that respect, we offer explicit models for these features and demonstrate the identification and generalization capabilities of the proposed models on a large database recorded in uncontrolled settings.

The outline of the paper is as follows. Section II elaborates on the related work, whereas Section III explains the experiment setup and dataset. Section IV models the distance, scalar product of velocity vectors, and scalar product of velocity and displacement vectors concerning coherent and non-coherent pairs. Finally, Section V presents the experimental results and Section VI summarizes our contributions.

II. RELATED WORK

Modeling and analysis of pedestrian motion has attracted considerable attention due to the vast application domain, such as simulation of crowd dynamics, evacuation processes and panic situations or planning of such facilities as stations and stadiums [1]. Lately the results of these studies have been applied to robotics to count and track groups of pedestrians [2], to model the collision avoidance and path planning of the robot in a way that is compatible with human behavior [3]–[5], and to design a companion robot which can walk along a group of people [6].

State-of-the-art methods in pedestrian detection and analysis benefit from the advances in sensor and tracking technologies and provide high accuracy localization of humans [7], [8]. Until recently the research on pedestrian behavior was focused on studying the behavior of the individual pedestrian, assuming that the interaction between different pedestrians was basically limited to collision avoiding. However, in the last years several studies have been conducted concerning the interactions of pedestrians walking in groups [9]–[12].

According to these works, pairs of pedestrians, who are socially related, walk usually side-by-side. Larger groups can assume more complex configurations, but according to [9] in any case at relatively low pedestrian densities they assume an abreast formation, i.e. each pedestrian can be considered as part of one or two side-by-side pairs.

Inspiring from the definition of 'group' given in [13], Ge et al. employ norms of velocity difference vectors of pedestrians to compute the locomotion similarity [11]. Subsequently, in order to detect pedestrian groups, a bottom-up hierarchical clustering scheme is carried out based on locomotion similarities. Sandıkcı et al. follow a similar strategy in detection of a group of pedestrians applying an agglomerative clustering on the trajectory similarity matrix, which integrates positional, velocity, and directional similarities [12]. Pellegrini et al. and Yamaguchi et al. consider detection of pedestrian groups as an enhancement on tracking in crowded environments where significant occlusion is observed [14], [15].

However, these methods of pedestrian behavior analysis focus on motion of pedestrians solely considering interaction to be possible only between humans. We, on the other hand, focus also on the interaction of humans with other objects, distinguishing mainly between the side-by-side behavior observed between pedestrians and the tandem configuration of pedestrians and pushcarts and the like. At this stage, we would like to emphasize the difference between 'group motion' and coherence. While in literature group motion is used in the sociological sense usually referring to the motion of a number of humans, who have a social connection and the same goal or intention [13], [16], coherence embraces group motion and expands its extent of interactions accounting for interaction of humans with various auxiliary instruments such as walking aids or pushcarts (see Figure 2).

III. EXPERIMENT SETUP AND PREPROCESSING

For investigating the inherent properties of coherence, an experiment is performed in the entrance hall of a shopping center between 15-16 PM on a weekday, where the area of tracking and classification is restricted to a rectangular region of 7.5 m × 8 m. Two sorts of recordings are gathered, namely a video recording of the scene, which is used for validation purposes, and range information, which is used for investigating the coherence relations numerically.

Based on the video recording, the ground truth relating the entity types and coherence qualities is established as in Figure 3. Here, non-coherent pairs involve single pedestrians, manual or electric wheelchairs and carts; while coherent pairs involve human-human, human-shopping cart and human-baby cart pairs. In this framework, side-by-side motion is presented usually by human-human pairs, whereas tandem motion is performed by human-cart pairs.
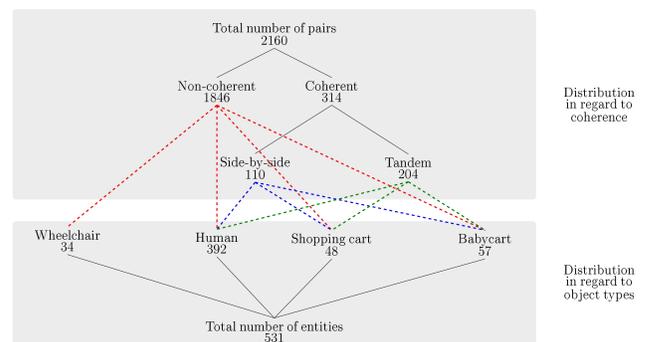

Fig. 3. Distribution of entities in regard to object type and coherence quality. The dashed lines represent possible coherence qualities that certain entity types may attain.

The range readings are collected by the four laser range finders positioned at the corners of the observation space, where the scanning plane is fixed at 87 cm above the ground to detect the humans and objects properly. The locations of the tracked entities are determined as the center of mass of the range samples. Here, note that the proposed method does not necessarily operate with range sensors and can be incorporated with different tracking methods as soon as satisfactory localization rates are provided. The velocity and displacement vectors are computed after a preprocessing applied on the estimated locations, which might be adjusted in accordance with the tracking method. Since range sensors do not provide exactly uniform readings over time and are prone to a certain amount of positioning error due to the clutter in the environment and operation accuracy, the set of the raw locations is smoothed through a convolution operation with a low pass filter, reducing the noise on positioning. Subsequently, in order to achieve uniform sampling over time, the output of the low pass filter is unified over an interval of $\Delta t = 100$ msec. Finally, the enhanced location values regarding entities e_i, $\{p_i[t]\}$, are obtained. The velocity vectors are computed by taking the difference between every other two values, $\vec{v}_i[t] = (p_i[t] - p_i[t - 2\Delta t]) / (2\Delta t)$, in order to enhance smoothing.

Fig. 2. Examples for non-coherent and coherent pairs. Non-coherent entities involve (a) single pedestrian and, (b) person on wheelchair, whereas coherent entities involve (c) side-by-side motion and tandem motion as person pushing (d) small baby cart, (e) big baby cart, or shopping cart.

IV. MODELING INDICATORS OF COHERENCE

From a coherence standpoint, the distance and relative velocity patterns are suggested to be the most relevant indicators. Particularly, the distance between pedestrians and scalar product of their velocities are suggested to highlight the distinctions between coherent and non-coherent pairs, whereas the scalar product of velocity and displacement is proposed to reflect the discrimination of side-by-side and tandem motion (see Figure 6). On that account, determining the coherence quality of a pair as coherent or non-coherent based on distance and scalar product of velocities and subsequently determining the particular nature as moving on the side or in tandem based on scalar product of velocity and displacement is a reasonable approach.

Our previous work follows this strategy and proposes a method for identification of coherence quality based on several simplifying assumptions imposed on these discriminative attributes [17]. Namely, for the sake of simplicity the entities involved in coherent motion are proposed to be characterized by perfectly aligned velocity vectors. In addition, for entities moving in an ideal side-by-side fashion, the velocity vectors are suggested to be perpendicular to the displacement vector, whereas for entities engaged in tandem motion, the velocity vectors are ideally parallel to the displacement vectors. Applying certain thresholds to account for the non-ideal nature of the behavior, coherence qualities of the entities are resolved with satisfactory success rates.

Although this simplistic method exploiting orthogonality and perfect alignment assumptions performs significantly well, an explicit model is necessary to improve its performance and make it flexible to adapt to different settings effectively. To that end, these features are investigated closely and a mathematical model is proposed for each of the relating probability density functions. The model parameters are resolved employing a least squares minimization together with a golden section search. The data is shuffled and a random set of observations composed of 10% of the entire data is chosen as a training set. The squared error between the proposed distribution and the training set is minimized to solve for the model parameters. Subsequently, the remaining data which constitutes 90% of observations is used as a test set. This procedure is repeated 100 times to investigate the stability of parameter estimation and performance in recognition of coherence quality.

A. Modeling distance

The observation environment is assumed to be completely uniform, i.e. the probability of visiting each point is equal,

$$P(p_m) = P(p_n), \forall p_m, p_n \in A, \quad (1)$$

where $P(p_m)$ denotes the probability of visiting point p_m and A stands for the observation area.

Let a point on the trajectory of an entity e_i at time t be represented by $p_i[t]$. Suppose that there are two entities e_i and e_j, which are observed simultaneously over an interval $[t_0, t_f]$. The distance between e_i and e_j at an intermediate time instant $t \in [t_0, t_f]$ is defined as the magnitude $\delta[t] = |\vec{\delta}[t]|$, where $\vec{\delta}[t]$ is the displacement from $p_i[t]$ to $p_j[t]$. For all possible combinations of i and j, the empirical distribution of δ_x-δ_y is found to be as in Figure 4-(a) and (b) for coherent and non-coherent pairs, respectively.

1) Modeling distance for coherent pairs: The components of the displacement vector $\vec{\delta}$ can be written as $\delta_x = \delta \cos(\alpha)$ and $\delta_y = \delta \sin(\alpha)$, where α stands for the argument of $\vec{\delta}$ according to an arbitrary reference frame (see Figure 5). Figure 4-(a) presents four prominent peaks around which δ_x and δ_y are roughly normally distributed. Namely,

$$\delta_x \sim \mathcal{N}(\nu \cos(\alpha), \sigma^2),$$
$$\delta_y \sim \mathcal{N}(\nu \sin(\alpha), \sigma^2). \quad (2)$$

Equation 2 suggests δ is distributed with a Rice distribution,

$$p(\delta|\nu, \sigma) = \frac{\delta}{\sigma^2} \exp\left(\frac{-\delta^2 - \nu^2}{2\sigma^2}\right) I_0\left(\frac{\delta \nu}{\sigma^2}\right), \quad (3)$$

where I_0 stands for the modified Bessel function of first kind with order 0 [18].

While the δ_x-δ_y distribution may present some peaks due to the presence of dominant motion directions, the distribution of δ according to Equation 3 is invariant to the orientation α. Thus, the distribution of δ, which involves several peaks, is still given by Equation 3. This result holds even in the absence of any prominent direction such that α is a uniformly distributed circular random variable.

The parameters ν and σ that describe the data best are found be 711 ± 23 mm and 127 ± 19 mm, respectively. The relative standard deviation is much smaller than 1, which indicates that the parameter estimation is robust against varying training sets, even though they represent a small fraction of the entire set of observations.
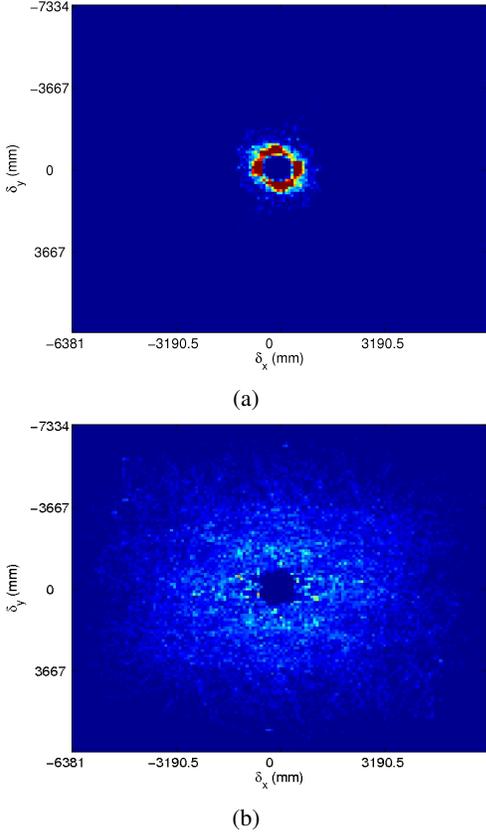
(a)

(b)

Fig. 4. The distribution of δ_x and δ_y between every possible pair presenting (a) coherent and (b) non-coherent motion.

2) Modeling distance for non-coherent pairs: If two entities e_i and e_j are non-coherent, then their relative locations at a particular instant t are independent. This assumption, together with Equation 1, allows us to model δ_x and δ_y in the following way. Suppose that the dimensions of the observation environment along $x-$ and $y-$axes are D, then

$$p(\delta_x) = \begin{cases} \frac{4}{D^2}\delta_x, & \text{if } 0 < \delta_x < \frac{D}{2}, \\ \frac{4}{D} - \frac{4}{D^2}\delta_x, & \text{if } \frac{D}{2} < \delta_x < D, \end{cases}$$

while the probability density function (pdf) concerning δ_y is computed in the same manner. Assuming that δ_x and δ_y are independent, the relating joint pdf is resolved as,

$$p(\delta) = \begin{cases} \frac{1}{D}2\delta\left(\delta^2 - 4\delta + \pi\right), & \text{if } 0 < \delta < D, \\ \frac{1}{D}2\delta[4\sqrt{\delta^2 - 1} - (\delta^2 + 2 - \pi) - \\ \quad 4\tan^{-1}\left(\sqrt{\delta^2 - 1}\right)] & \text{if } D < \delta < D\sqrt{2}. \end{cases}$$

This distribution describes δ regarding non-coherent pedestrians in a large environment, namely $D \gg c$, where $c \approx 400$ mm stands for the physical size of the human body. However, it does not consider the constraint imposed by the physical dimensions of pedestrians, which represent a cutoff below which δ cannot assume values. To account for this cutoff, δ is substituted with $\delta' = \delta - c$, and $p(\delta)$ is renormalized by replacing D with $D' = D - c/\sqrt{2}$. Hence this distribution does not need to be calibrated since it depends only on the geometry of the observation area.
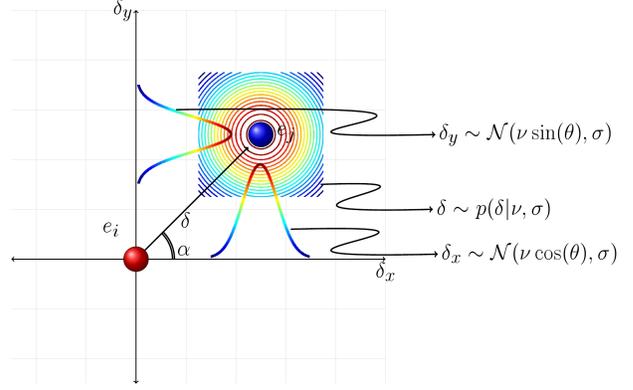
Fig. 5. The distribution of δ_x, δ_y and δ.

B. Modeling scalar product of velocities

In explicit terms, the scalar product of the velocity vectors $\vec{v}_i$ and $\vec{v}_j$ is,

$$\frac{\vec{v}_i \cdot \vec{v}_j}{|\vec{v}_i||\vec{v}_j|} = \cos\left(\theta_{ij}\right), \tag{4}$$

where θ denotes the angle between these vectors (see Figure 6). In order to model this scalar product, it is proposed first to obtain a model of θ and then apply it to Equation 4.

Coherent pairs are expected to have the direction of velocity vectors aligned most of the times, whereas non-coherent pairs do not present any correlation of direction. This suggests that the expected value of θ is 0 for both coherent and non-coherent pairs. If θ were a continuous variable over $(-\infty, \infty)$, such a behavior could be approximated with $\mathcal{N}(0, \sigma_\theta^2)$. However, θ is defined over the range $[-\pi, \pi]$ and, thus, it cannot be modeled as a standard normal distribution. Hence, the principles of directional statistics are invoked and the behavior of θ is modeled as a *von Mises* distribution [19], which is the circular analogue of the Gaussian distribution. The explicit form of the von Mises distribution is,

$$p(\theta|\mu, \kappa) = \frac{\exp(\kappa\cos(\theta - \mu))}{2\pi I_0(\kappa)}, \tag{5}$$

where μ denotes the mean value and κ is the analogous of $1/\sigma^2$ of the normal distribution. In other words, for $\kappa \gg 1$ the distribution is very localized around μ, while for $\kappa \to 0$ the distribution is uniform. Note that the θ distribution relating coherent and non-coherent pairs is described using the same function given by Equation 5, where the parameter κ enables modeling of different behavior.

By replacing $\mu = 0$ in Equation 5, making the change of variables $z = \cos(\theta)$ and arranging the signs of the trigonometric terms for corresponding ranges of z, we get,

$$p(z|\kappa) = \frac{\exp(\kappa z)}{\pi I_0(\kappa)}\frac{1}{\sqrt{1 - z^2}}.$$

The resolved values of κ are found to be 18.35 ± 0.78 for coherent pairs and 0.06 ± 0.08 for non-coherent pairs. These indicate that the distribution of θ is localized around 0 for coherent pairs, while it is almost uniform for non-coherent pairs, as we expected.
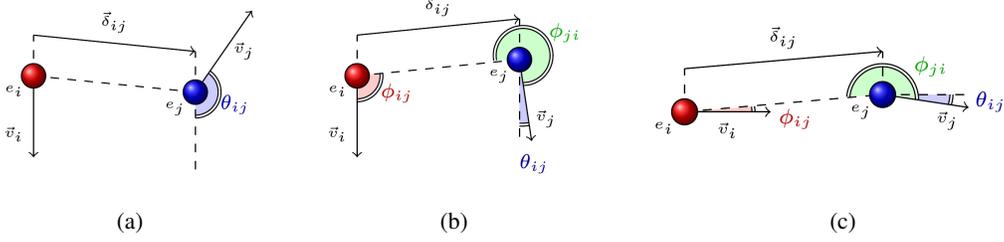
Fig. 6. (a) A non-coherent pair, (b) a coherent pair engaged in side-by-side motion, and (c) a coherent pair engaged in tandem motion,. Here $\vec{v}_i$ and $\vec{v}_j$ are the velocity vectors and $\vec{\delta}_{ij}$ stands for the displacement from e_i to e_j.

C. Modeling scalar product of velocity and displacement

The scalar product of velocity and displacement vectors concerning entities e_i and e_j is given by,

$$\frac{\vec{v}_i \cdot \vec{\delta}_{ij}}{|\vec{v}_i| \, |\vec{\delta}_{ij}|} = \cos(\phi_{ij}),$$

where ϕ_{ij} is the angle between the velocity vector of e_i, $\vec{v}_i$, and the displacement vector from e_i to e_j, $\vec{\delta}_{ij}$ (see Figure 6). Here $\cos(\phi)$ is proposed to enable discrimination of the particular nature of coherence as side-by-side or in tandem. Thereby, only the entities which are known to be coherent are examined and two models are proposed for the pdf of $\cos(\phi)$ for entities moving side-by-side and in tandem. To obtain a model for the pdf of $\cos(\phi)$, a model similar to that of θ is used for modeling ϕ and the distribution of $\cos(\phi)$ is derived therefrom for tandem and side-by-side motion.

The mean value μ appearing in the von Mises distribution relating ϕ takes values other than 0 unlike the model of scalar product of velocity vectors. For coherent pairs moving side-by-side, ϕ is either $\pi/2$ or $3\pi/2$ depending on whether the entity of interest is moving on the right or left. Assuming that for each entity engaged in side-by-side motion and moving on the right, there is another entity moving on its left, the probability that $\phi = \pi/2$ is equal to the probability that $\phi = 3\pi/2$. Thus, ϕ comes from an equally weighted linear combination of two von Mises distributions,

$$p(\phi|\kappa) = \frac{1}{2} \frac{\exp(\kappa \cos(\phi - \pi/2))}{2\pi I_0(\kappa)} + \frac{1}{2} \frac{\exp(\kappa \cos(\phi - 3\pi/2))}{2\pi I_0(\kappa)}. \tag{6}$$

By making the change of variables $z = \cos(\phi)$ and simplifying the trigonometric terms, the pdf of $\cos(\phi)$ relating side-by-side moving entities is obtained as,

$$p(z|\kappa) = \frac{\cosh\left(\kappa\sqrt{1 - z^2}\right)}{\pi I_0(\kappa)\sqrt{1 - z^2}}. \tag{7}$$

On the other hand, in tandem motion, μ is 0 for the entity moving at the rear position and π for the entity moving at the front position (see Figure 6). Applying this to the general form of the distribution presented in Equation 5, and taking an equally weighted linear combination of two distributions as in Equation 6, the model of z for tandem motion is,

$$p(z|\kappa) = \frac{\cosh(\kappa z)}{\pi I_0(\kappa)\sqrt{1 - z^2}}. \tag{8}$$

The optimal values of κ are found as 1.73 ± 0.76 and 18.90 ± 1.62 for side-by-side moving entities and entities engaged in tandem motion, respectively, showing that the ϕ distribution is much more localized for tandem pairs. This indicates that tandem motion is more structured than side-by-side motion.

V. EXPERIMENTAL RESULTS

This section compares the proposed models of the indicators of coherence with the observed distributions. A detailed discussion on the performance of description of the identifying differences between the behaviors by the models is supported by a log-likelihood based quantification. Coherence characteristics of the pairs in the test set, which contains 90% of all pairs, are estimated by utilizing the models of δ, $\cos(\theta)$ and $\cos(\phi)$, which are calibrated according to a non-overlapping training set, that contains the remaining 10% of all the pairs. To that end, δ, $\cos(\theta)$ and $\cos(\phi)$ relating a query pair are replaced in the proposed models and the coherence quality of the pairs is estimated as the one with the higher likelihood. Subsequently, the rate of correct detections is presented in a tabular form.

A. Performance in modeling of distance

The observed distribution and the proposed model of δ for coherent and non-coherent pairs are illustrated in Figure 7. This figure ascertains that both models are able to grasp the peaks efficiently.
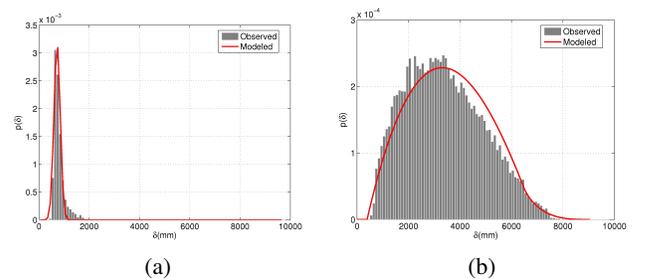

Fig. 7. The empirical distribution and the proposed model for the distribution of δ for (a) coherent and (b) non-coherent pairs.

The performance of the proposed models is verified by comparing likelihoods. The first row of Table I indicates that on the average 99.24% of non-coherent pairs and 89.36% of

TABLE I

PERFORMANCE IN DISTINCTION OF COHERENCE CHARACTERISTICS USING LOG-LIKELIHOOD OF THE PROPOSED MODELS.

	Non-coherent (%)	Coherent (%)		Total (%)
δ	99.24 ± 0.38	89.36 ± 2.29		98.92
$\cos(\theta)$	91.64 ± 0.45	83.64 ± 1.55		91.52
Yücel et al. [17]	88.08	89.80		88.33
		Tandem	Side-by-side	
$\cos(\phi)$	-	95.09 ± 2.05	97.54 ± 1.10	96.68
Yücel et al. [17]		99.01	64.54	86.94

coherent pairs are identified correctly based on the observations of distance. Moreover, the standard deviations of the detection rates are concluded to be insignificant verifying that the method is robust against varying training sets.

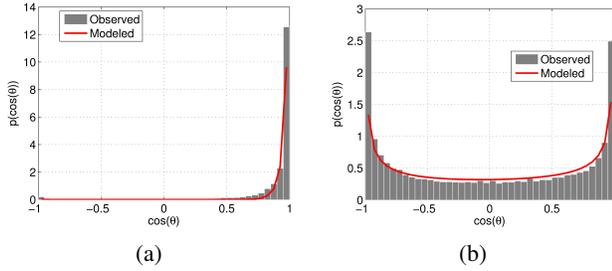### B. Performance in modeling of scalar product of velocities

Fig. 8. The empirical distribution and proposed model of $\cos(\theta)$ for (a) coherent and (b) non-coherent pairs.

The empirical pdfs and corresponding models of $p(\cos(\theta))$ are given in Figure 8. This figure supports our claim that $\cos(\theta)$ has a a distinctive behavior in terms of coherence quality. Moreover, the developed models are able to grasp the different patterns. Table I presents the average rate of detection based on likelihoods of $\cos(\theta)$ as 91.64% for non-coherent pairs and 83.64% for coherent pairs.

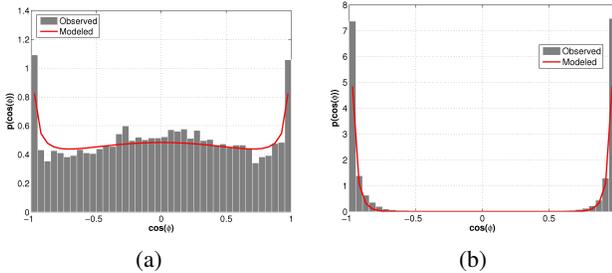### C. Performance in modeling of scalar product of velocity and displacement vectors

Fig. 9. The empirical distribution and proposed model of $\cos(\phi)$ for (a) side-by-side moving pairs, and (b) entities engaged in tandem motion.

As the resolved κ values are replaced in Equations 7 and 8, the models shown in Figure 9 are obtained. A comparison

of the empirical distributions of $\cos(\phi)$ indicates that similar to the previous models, the peaks are captured successfully in all cases. By applying these models to the observations of $\cos(\phi)$ coming from coherent pairs, 97.54% of entities moving in a side-by-side fashion and 95.09% of the entities moving in tandem are identified correctly.

D. Overlapping mis-detections

We investigate the overlapping mis-detections occurring in identification of coherence based on distance and scalar product of velocities solely. Figure 10-(a) illustrates that 4.89% of coherent pairs are mis-detected based on δ while they are correctly detected based on $\cos(\theta)$. Moreover, 10.61% of coherent pairs are mis-detected using the log-likelihood of $\cos(\theta)$ but they are correctly detected based on δ. This suggests that when the inferences coming from the two indicators are incorporated, there is a potential to correct most of them. Similarly, Figure 10-(b), presents the overlapping mis-detections concerning the non-coherent pairs. For this case, 0.17% of the pairs are mis-detected by both inferences regarding δ and $\cos(\theta)$. Apart from these common mis-detections, most of the erroneous cases have the potential to be corrected. Thereby, it is suggested that a combination of δ and $\cos(\theta)$ enables enhancement in resolution of coherence characteristics.
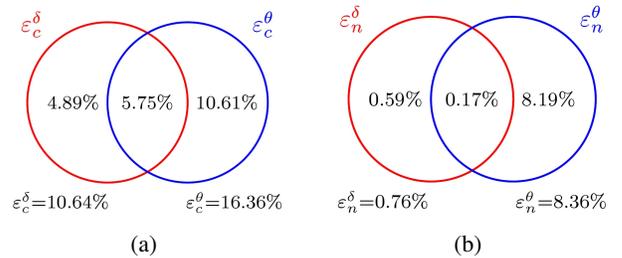

Fig. 10. Overlapping mis-detections in identification of (a) coherent pairs and (b) non-coherent pairs based on δ and θ separately. ε_c^δ and ε_n^δ denote the rate of mis-detections in detection of coherent and non-coherent pairs, respectively, using δ. Similarly, ε_c^θ and ε_n^θ are for θ.

E. Comparison

Table I presents the performance rates of the proposed models together with those of the method described in [17].

It is evident that the model of δ improves detection of both coherent and non-coherent entities, whereas the model of $\cos(\theta)$ improves the detection rate of non-coherent pairs and results in a small degradation of performance in identification of coherent pairs. Nonetheless, in the overall detection rate both models outperform [17]. Namely, the recognition performance using δ increases roughly by 10% from 88.83% to 98.92%, while the improvement in recognition using $\cos(\theta)$ is approximately 3%. Regarding the resolution of the particular configuration of coherence relation as tandem or side-by-side, our method introduces a significant improvement over that of [17] in detection of entities engaged in side-by-side motion. Moreover, it is clear from Table I that [17] has a positive bias for recognizing the coherence relation often as tandem, which results in a very low recognition rate of 64.54% for side-by-side motion. However, our method achieves comparable detection accuracies eliminating the bias. In addition, the proposed method outperforms [17] with a 10% improvement in terms of overall detection rate, namely from 86.94% to 96.68%. Furthermore, one should note that [17] depends on the empirical data in the selection of several thresholds. In that respect, our method offers a flexibility due to the automatic estimation of the model parameters through a training with a small dataset.

Moreover, our approach enables the computation of the likelihood that a query pairs comes from a particular distribution, employing the modeled probability density functions of δ, $\cos(\theta)$ and $\cos(\phi)$, unlike [17], which provides only a binary decision for each of these features. This implies that our method allows a quantitative comparison of the estimations enabling the correction of non-overlapping mis-detections as suggested in Section V-D, which is impossible to apply in the framework of [17].

F. Generalization capabilities

The proposed method is flexible in many respects. First of all, various tracking techniques such as vision- or range-based ones can be incorporated with our approach as soon as they provide satisfactory location estimation accuracy. In addition, provided that the condition of uniformity given in Equation 1 is satisfied, our method is not specific to the environment and thereby has the possibility of being adapted onto various domains. Nevertheless, the parameters of the models are expected to depend on features of the environment and/or its users to a certain extent. This flexibility of the method enables characterization of these different features through the values of its specifying parameters. Moreover, the parameters of the models are possible to calibrate from a small training set. Namely, ν, σ and the κ's are demonstrated to be estimated from only 10% of the whole samples effectively, which leads to significantly high performance rates in identification of the coherence.

VI. CONCLUSION

This study examines the joint motion of human-human pairs as well as human-object pairs. Identification of such entities with considerable correlation in trajectory and velocity is a crucial point in improvement of pedestrian tracking in crowded environments and under occlusion as well as motion planning and collusion avoidance. In order to get a better insight for coherence, distance, scalar product of velocity vectors and scalar product of velocity and displacement vectors are investigated and a mathematical model is developed for each. The proposed approach is shown to grasp the characterizing features of different patterns of coherent motion efficiently. Moreover, the models are flexible to be adapted onto different environments and easy to calibrate even with a small set. An improvement in discrimination of coherent and non-coherent pairs can be attained by the investigation of overlapping mis-detections.

REFERENCES

[1] D. Helbing and A. Johansson, *Pedestrian, crowd, and evacuation dynamics*. New York: Springer, 2009, vol. 16, pp. 6476–6495.
[2] B. Lau, K. Arras, and W. Burgard, "Multi-model hypothesis group tracking and group size estimation," *International Journal of Social Robotics*, vol. 2, no. 1, pp. 19–30, 2010.
[3] Y. Tamura, T. Fukuzawa, and H. Asama, "Smooth collision avoidance in human-robot coexisting environment," in *IROS*, 2010, pp. 3887–3892.
[4] P. Henry, C. Vollmer, B. Ferris, and D. Fox, "Learning to navigate through crowded environments," in *ICRA*, 2010, pp. 981–986.
[5] T. Bandyopadhyay, N. Rong, M. Ang, D. Hsu, and W. Lee, "Motion planning for people tracking in uncertain and dynamic environments," in *Workshop on people detection and tracking, ICRA*, 2009.
[6] A. Carballo, A. Ohya, and S. Yuta, "Reliable people detection using range and intensity data from multiple layers of laser range finders on a mobile robot," *International Journal of Social Robotics*, pp. 1–20, 2011.
[7] O. Mozos, R. Kurazume, and T. Hasegawa, "Multi-part people detection using 2d range data," *International Journal of Social Robotics*, vol. 2, no. 1, pp. 31–40, 2010.
[8] K. Arras, O. Mozos, and W. Burgard, "Using boosted features for the detection of people in 2d range data," in *ICRA*. IEEE, 2007, pp. 3402–3407.
[9] M. Moussaïd, N. Perozo, S. Garnier, D. Helbing, and G. Theraulaz, "The walking behaviour of pedestrian social groups and its impact on crowd dynamics," *PLoS ONE*, vol. 5, no. 4, p. e10047, 2010.
[10] M. Costa, "Interpersonal distances in group walking," *Journal of Nonverbal Behavior*, vol. 34, no. 1, pp. 15–26, 2010.
[11] W. Ge, R. Collins, and B. Ruback, "Automatically detecting the small group structure of a crowd," in *Workshop on Applications of Computer Vision*, 2009, pp. 1–8.
[12] S. Sandıkcı, S. Zinger, and P. de With, "Detection of human groups in videos," *Advances Concepts for Intelligent Vision Systems*, pp. 507–518, 2011.
[13] C. McPhail and R. Wohlstein, "Using film to analyze pedestrian behavior," *Sociological Methods & Research*, vol. 10, no. 3, p. 347, 1982.
[14] S. Pellegrini, A. Ess, and L. Van Gool, "Improving data association by joint modeling of pedestrian trajectories and groupings," *ECCV*, pp. 452–465, 2010.
[15] K. Yamaguchi, A. C. Berg, L. E. Ortiz, and T. L. Berg, "Who are you with and where are you going?" in *CVPR*, 2011, pp. 1345–1352.
[16] M. Moussaïd, D. Helbing, S. Garnier, A. Johansson, M. Combe, and G. Theraulaz, "Experimental study of the behavioural mechanisms underlying self-organization in human crowds," *Proc. the Royal Society B: Biological Sciences*, vol. 276, no. 1668, pp. 2755–2762, 2009.
[17] Z. Yucel, I. Ikeda, T. Miyashita, and N. Hagita, "Identification of Mobile Entities Based on Trajectory and Shape Information," in *IROS*, 2011, pp. 1–6.
[18] M. Abramowitz and I. Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. Dover publications, 1964, vol. 55, no. 1972.
[19] K. V. Mardia and P. E. Jupp, *Directional statistics*. John Wiley & Sons Inc, 2000.