International Journal of Social Robotics manuscript No.
(will be inserted by the editor)

Multi-Model Hypothesis Group Tracking and Group Size Estimation

Boris Lau · Kai O. Arras · Wolfram Burgard

Received: data / Accepted: date

Abstract People tracking is essential for robots that are supposed to interact with people. The majority of approaches track humans in the vicinity of the robot independently. However, people typically form groups that split and merge. These group formation processes reflect social relations and interactions that we seek to recognize in this paper. To this end, we pose the group tracking problem as a recursive multi-hypothesis model selection problem in which we hypothesize over both, the partitioning of tracks into groups (models) and the association of observations to tracks (assignments). Model hypotheses that include split, merge, and continuation events are first generated in a data-driven manner and then validated by means of the assignment probabilities conditioned on the respective model. Observations are found by clustering points from a laser range finder and associated to existing group tracks using the minimum average Hausdorff distance. We further propose a method to estimate the number of people in the individual groups. Experiments with a mobile robot demonstrate that the approach is able to accurately recover social grouping of people with respect to the ground truth. The results also show that tracking groups is clearly more efficient than tracking people separately. Our system runs in real-time on a typical desktop computer.

Boris Lau · Wolfram Burgard
Autonomous Intelligent Systems Group,
Department of Computer Science, University of Freiburg, Germany,
Email: {lau,burgard}@informatik.uni-freiburg.de.

Kai O. Arras
Social Robotics Laboratory,
Department of Computer Science, University of Freiburg, Germany,
Email: arras@informatik.uni-freiburg.de.

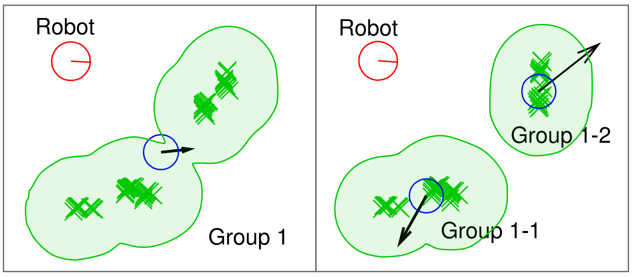

Fig. 1 Tracking groups of people with a mobile robot. Groups are shown by their position (blue), velocity (black), the associated laser points (green), and a contour for visualization. In the two frames, a group of four people splits up into two groups with two people each.

1 Introduction

The ability of robots to keep track of people in their surrounding is fundamental for a wide range of applications including personal and service robots, intelligent cars, crowd control, and surveillance. People are social beings and as such they form groups, interact with each other, merge to larger groups, or separate from groups. Tracking individual people in these formation processes can be hard due to the high chance of occlusion and the large extent of data association ambiguity. This causes the space of possible associations to become huge and the number of assignment histories to quickly become intractable. Further, for many applications, knowledge about groups can be sufficient as the task does not require to know the state of every person. In such situations, tracking groups that consist of multiple people is more efficient. Additionally, it reveals semantic information about activities and social relations of people.

This paper focuses on group tracking in populated environments with the goal to track a large number of people in real-time. The approach attempts to maintain the state of groups of people over time, considering possible splits and merges as shown in Fig. 1. For our experiments we use a mobile robot equipped with a laser range finder, but our method should be applicable to data from other sensors as well.

In most of the related work on laser-based people tracking, tracks correspond to individual people [1–5]. In Taylor *et al.* [6] and Arras *et al.* [7], tracks represent the state of legs which are fused to people tracks in a later stage. Khan *et al.* [8] proposed an MCMC-based tracker that is able to deal with non-unique assignments, i.e., measurements that originate from multiple tracks, and multiple measurements that originate from the same track. Actual tracking of groups using laser range data was, to our knowledge, first addressed by Mucientes *et al.* [9]. Most research in group tracking was carried out in the vision community [10–12]. Gennari *et al.* [11] and Bose *et al.* [12] both address the problem of target fragmentation (splits) and grouping (merges). They do not integrate data association decisions over time – a key property of the Multi-Hypothesis Tracking (MHT) approach, initially presented by Reid [13] and later extended by Cox *et al.* [14]. The approach belongs to the most general data association techniques as it produces joint compatible assignments, integrates them over time, and is able to deal with track creation, matching, occlusion, and deletion.

The works closest to this paper are Mucientes *et al.* [9] and Joo *et al.* [15]. Both address the problem of group tracking using an MHT approach. Mucientes *et al.* employ two separate MHTs, one for the regular association problem between observations and tracks, and a second stage MHT that hypothesizes over group merges. However, people tracks are not replaced by group tracks, hence there is no gain in efficiency. The main benefit of that approach is the additional semantic information about the formation of groups.

Joo *et al.* [15] present a vision-based group tracker using a single MHT to create hypotheses of group splits and merges and observation-to-track assignments. They develop a variant of Murty's algorithm [16] that generates the k-best *non-unique* assignments which enables them to make multiple assignments between observations and tracks, thereby describing target splits and merges. However, the method only produces an approximation of the optimal k-best solutions since the posterior hypothesis probabilities depend on the number of splits, which, at the time when the k-best assignments are being generated, is unknown. In our approach, the split, merge and continuation events are given by the model *before* computing the assignment probabilities, and therefore, our k-best solutions are optimal.

In this paper we propose a tracking system for groups of people using an extended MHT approach to hypothesize over both, the group formation process (models) and the association of observations to tracks (assignments). Each model, defined to be a particular partitioning of tracks into groups, creates a new tree branch with its own assignment problem. As a further contribution we propose a group representation that includes the shape of the group, and we show how this representation is updated in each step of the tracking cycle. This extends previous approaches to group tracking where groups are assumed to have Gaussian shapes only [11,9]. The group tracker proposed in this paper also estimates the number of people in groups and employs a labeling system to represent the history of group interactions, both of which extend the approach presented in our previous work [17].

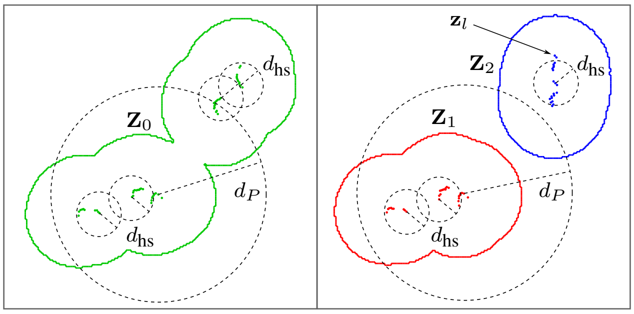

Fig. 2 Illustration of the detection step. *Left:* One group is detected since all shortest links between the measured points z_l are smaller than the single-linkage clustering threshold d_P. *Right:* Two groups are found as the shortest link between their points exceeds d_P. For group size estimation, the number of human-sized blobs in a group is determined by applying the same clustering procedure with threshold d_{hs}.

Finally, we use the psychologically motivated *proxemics* theory introduced by Hall [18] for the definition of a group. The theory relates social relation and body spacing during social interaction and proposes thresholds that separate the intimate, personal, social, and public space around people.

This paper is structured as follows: the following section describes the extraction of groups of people from laser range data. Section 3 introduces the definition of groups and group tracks. Section 4 briefly describes the cycle of our Kalman filter-based tracker. Section 5 explains the data-driven generation of models and how their probabilities are computed. Whereas Section 6 presents the multi-model MHT formulation and derives expressions for the hypothesis probabilities, Section 7 describes the experimental results.

2 Group Detection in Range Data

Detecting people in range data has been approached with motion and shape features [1–5,9] as well as with a learned classifier using boosted features [19]. However, these systems were designed (or trained) to extract single people. In the case of densely populated environments, groups of people typically produce large blobs in which individuals are hard to recognize. We therefore pursue the approach of background subtraction and clustering. Given a previously learned model (a map of the environment for mobile platforms), the background is subtracted from the scans and the remaining points are passed to the clustering algorithm. This approach is also able to detect standing people as opposed to the work of Mucientes *et al.* [9] which relies on motion features. Note that the detection method is not critical to the system and could also be replaced by map-free approaches that employ appearance information, motion features, or other filtering techniques.

Concretely, a laser scanner generates measurements consisting of bearing and range values. The measurements are transformed into Cartesian coordinates $\mathbf{z}_l = (x_l, y_l)^T$ and grouped using *single linkage clustering* [20] with a distance threshold d_P. The outcome is a set of clusters $\mathcal{Z}_i$ making up the current observation set $Z(k) = \{\mathcal{Z}_i \,|\, i = 1, \ldots, N_{\mathcal{Z}}\}$. Each cluster $\mathcal{Z}_i$ is a complete set of measurements $\mathbf{z}_l$ that fulfills the cluster condition, i.e., two clusters are joined if the distance between their closest points is smaller than d_P. A similar concept, using a connected components formulation, has been used by Gennari and Hager [11]. The clusters then contain range readings that can correspond to single legs, individual people, or groups of people, depending on the cluster distance d_P.

Even though tracking of individuals in groups is not feasible due to frequent occlusions, the number of detected individuals in a group correlates with the true number of people in a group. As an observation of the group size, we therefore take the number of human-sized clusters $n_{\text{hs}}(\mathcal{Z}_i)$ found in an observation cluster $\mathcal{Z}_i$. We determine this by counting the clusters after reapplying single linkage clustering to the points in $\mathcal{Z}_i$ with an appropriate distance threshold d_{hs}, with $d_{\text{hs}} < d_P$.

An example for the clustering is given in Fig. 2. On the left, all links are shorter than d_P so that the measurements are grouped into one cluster $\mathcal{Z}_0$ that contains four human-sized clusters. On the right, the shortest distance between the two groups exceeds d_P so that they are kept as two clusters, $\mathcal{Z}_1$ and $\mathcal{Z}_2$. The two people in $\mathcal{Z}_2$ are counted as only one human-sized cluster.

3 Group Definition and Group Tracks

This section defines the concept of a group, describes the initialization of group tracks and derives the probabilities of group-to-observation assignments and group-to-group assignments.

What makes a collection of people a *group* is a highly complex question in general, which involves social relations among subjects that are difficult to measure. A concept related to this question is the proxemics theory introduced by Hall [18] who found from a series of psychological experiments that social relations among people are reliably correlated with physical distance during interaction. This finding allows us to infer group affiliations by means of body spacing information available in the range data. The distance d_P thereby becomes a threshold with a meaning in the context of group formation.

3.1 Representation of Groups

Concretely we represent a group as a tuple $G = \langle \mathbf{x}, C, \mathcal{P}, \mathcal{L} \rangle$ with $\mathbf{x}$ as the track state, C the state covariance matrix, $\mathcal{P}$ the set of contour points that belong to G, and $\mathcal{L}$ the set of identification labels. The track state vector $\mathbf{x} = (x,\, y,\, \dot{x},\, \dot{y}, n)^T$ is composed of the position $(x,\, y)^T$, the velocities $(\dot{x},\, \dot{y})^T$, and n, the number of people in the group.

The points $\mathbf{x}_{\mathcal{P}_l} \in \mathcal{P}$ are an approximation of the current shape or spatial extension of the group. Shape information will be used for data association under the assumption of *instantaneous rigidity*. That is, a group is assumed to be a rigid object over the duration of a time step Δt, and consequently, all points in $\mathcal{P}$ move coherently with the estimated group state $\mathbf{x}$. The points $\mathbf{x}_{\mathcal{P}_l}$ are represented relative to the state $\mathbf{x}$.

The label set $\mathcal{L}$ contains identification labels that are associated with the group. These labels explicitly represent the history of track interactions, which can be of high interest for certain applications, e.g., to determine which people belong together.

3.2 Initialization of Group Tracks

If the tracker creates a new group track G_j from an observation cluster $\mathcal{Z}_i$ in time step k, the positional components $(x_j, y_j)^T$ of track state $\mathbf{x}_j(k|k)$ are initialized with the centroid position of the measurement cluster. The contour points $\mathcal{P}_j$ are the points in $\mathcal{Z}_i$ represented relative to the centroid (omitting the time index $(k|k)$ for readability):

$$\begin{pmatrix} x_j \\ y_j \end{pmatrix} := \bar{\mathbf{z}}_i = \frac{1}{|\mathcal{Z}_i|} \sum_{z_l \in \mathcal{Z}_i} z_l \ , \quad \mathcal{P}_j := \bigcup_{z_l \in \mathcal{Z}_i} z_l - \bar{\mathbf{z}}_i \ . \quad (1)$$

The unobserved velocity components $(\dot{x}_j, \dot{y}_j)^T$ of $\mathbf{x}$ are set to zero, the size estimate is set to the number of human-sized blobs in the measurement cluster, $n_j := n_{\text{hs}}(\mathcal{Z}_i)$, and the label set is assigned a unique number as its only element, e.g., $\mathcal{L}_j := \{0\}$ for the first group after starting up the tracker. The initial state covariance is given by $C_j = C_0$, where C_0 is a diagonal matrix with $\left(\sigma_x{}^2, \sigma_y{}^2, \sigma_{\dot{x}}{}^2, \sigma_{\dot{y}}{}^2, \sigma_n{}^2\right)$ being the elements on the main diagonal. To account for the unknown components in the initial state vector, high uncertainty values are used for the corresponding entries in the initial state covariance matrix.

3.3 Motion Model for Group Tracks

To track groups over time, the state $x(k|k)$ and state covariance $C(k|k)$ of each group track in time step k are predicted into the next time step using a motion model. The predictions are denoted as $x(k{+}1|x)$ and $C(k{+}1|k)$, respectively. For tracks that are *continued*, i.e., no splits or merges take place from one frame to the next, we assume constant velocity for the centroid of the group, and a constant number of people in the group. Using a linear Kalman filter we get $\mathbf{x}(k{+}1|k) = A\,\mathbf{x}(k|k)$ and $C(k{+}1|k) = A\,C(k|k)\,A^T + Q$ for the state prediction. The state transition matrix A and the process noise covariance matrix Q are given by

$$A = \begin{pmatrix} 1 & 0 & \Delta t & 0 & 0 \\ 0 & 1 & 0 & \Delta t & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \; Q = \begin{pmatrix} {\epsilon_x}^2 & 0 & 0 & 0 & 0 \\ 0 & {\epsilon_y}^2 & 0 & 0 & 0 \\ 0 & 0 & {\epsilon_{\dot{x}}}^2 & 0 & 0 \\ 0 & 0 & 0 & {\epsilon_{\dot{y}}}^2 & 0 \\ 0 & 0 & 0 & 0 & {\epsilon_n}^2 \end{pmatrix}.$$

The entries of Q reflect the acceleration capabilities of a typical human. The noise for the number of people in the group, controlled by ϵ_n, accounts for people joining or leaving the group without being noticed. The actual noise values used in our experiments are given in Sect. 7.

As mentioned before, we assume instantaneous rigidity for the shape of a group. Since the points in $\mathcal{P}$ are relative to the moving centroid, the point set remains unchanged, and $\mathcal{P}(k+1|k) = \mathcal{P}(k|k)$.

If two observations can be associated with a group track G, i.e., they both fall into the validation gate of G, the tracker can consider to *split* the track into two new tracks according to an interaction model (see Sect. 5). Since the actual partitioning in the split is unknown at this stage, two new predicted group tracks G_1 and G_2 are created by duplicating the predicted state and covariance of G. The same applies for the point set $\mathcal{P}$ and the label set $\mathcal{L}$. To make the label sets unique, we attach different indices to the label, e.g., a group with label set $\{0\}$ would split up into two groups with label sets $\{0-0\}$ and $\{0-1\}$. Again, the component of the state that represents the number of people in the group, n, is treated differently: the sum of people in the resulting groups must be equal to the original number of people. However, the actual partitioning is not known in the prediction step. Therefore, we use $n_1 = n_2 = n/2$, and reinitialize the state covariances of the new split tracks with C_0.

If the tracker considers to *merge* two group tracks G_i and G_j according to a track interaction model, the track prediction has to be computed accordingly. The predicted set of contour points of the merged group is the union of the two former point sets, $\mathcal{P}_{ij} = \mathcal{P}_i \cup \mathcal{P}_j$. The track states $\mathbf{x}_i$ and $\mathbf{x}_j$ of the merging group track represent the position and velocity of the centroids of the groups. Thus, the state of the merged track, $\mathbf{x}_{ij}$, is computed as the weighted mean of the original track states, using the number of points in the merging sets $\mathcal{P}_i$ and $\mathcal{P}_j$ as weights. The tracks before the merge are assumed to be independent. According to the summation and scaling laws for covariances, the covariance matrix of the merging track is the weighted mean of the original covariances with squared weights,

$$\mathbf{x}_{ij} = w_i \cdot \mathbf{x}_i + w_j \cdot \mathbf{x}_j \tag{2}$$

$$C_{ij} = {w_i}^2 \cdot C_i + {w_j}^2 \cdot C_j\,, \tag{3}$$

where $w_i = |\mathcal{P}_i|/|\mathcal{P}_{ij}|$ and $w_j = |\mathcal{P}_j|/|\mathcal{P}_{ij}|$. Note that this applies only for the first four components of $\mathbf{x}_{ij}$ and the upper-left 4×4 block of C_{ij}. The fifth component, namely the group size n_{ij}, is excluded, since the number of people in the merging groups naturally add up to $n_{ij} := n_i + n_j$. Consequently, the corresponding uncertainty values are summed up as well. Finally, the label set of the new group is the union of the label sets of the original tracks, $\mathcal{L}_{ij} = \mathcal{L}_i \cup \mathcal{L}_j$. To remove redundant labels, an optional pruning can be done in this step: whenever all tracks that resulted from a split have merged again, the additional indices added in the split step can be removed, e.g., when the groups with labels $\{0-0\}$ and $\{0-1\}$ merge, they can be labeled $\{0\}$ again. Although this can remove split and merge events from the history represented by the labeling, it keeps the semantic information consistent.

3.4 Group-to-Observation Assignment Probability

For data association we need to calculate the probability that an observed cluster $\mathcal{Z}_i$ belongs to a predicted group $G_j = \langle \mathbf{x}_j(k+1|k),\, C_j(k+1|k),\, \mathcal{P}_j \rangle$. Therefore, we are looking for a distance function $d(\mathcal{Z}_i, G_j)$ that, unlike the Mahalanobis distance used by Mucientes *et al.* [9], accounts for the shape of the observation cluster $\mathcal{Z}_i$ and the contour $\mathcal{P}_j$ of the group, rather than just for their centroids. To this end, we use a variant of the Hausdorff distance. As the regular Hausdorff distance is the *longest* distance between points on two contours, it tends to be too sensitive to large variations in depth that can occur in range data. This motivates the use of the minimum average Hausdorff distance [21] that computes the minimum of the averaged distances between contour points as

$$d_{\text{HD}}(\mathcal{Z}_i, G_j) = \min\{d(\mathcal{Z}_i, \mathcal{P}_j), d(\mathcal{P}_j, \mathcal{Z}_i)\}\,, \tag{4}$$

where $d(\mathcal{Z}_i, \mathcal{P}_j)$ is the directed average Hausdorff distance from $\mathcal{Z}_i$ to $\mathcal{P}_j$,

$$d(\mathcal{Z}_i, \mathcal{P}_j) = \frac{1}{|\mathcal{Z}_i|} \sum_{\mathbf{z}_l \in \mathcal{Z}_i} \min_{\mathbf{x}_{\mathcal{P}_j} \in \mathcal{P}_j} \{D(\nu_{lj},\, S_{lj})\}\,. \tag{5}$$

Since we deal with uncertain entities, we calculate the distance $d(\mathcal{Z}_i, \mathcal{P}_j)$ using the Mahalanobis distance

$$D(\nu_{lj}, S_{lj}) = \sqrt{{\nu_{lj}}^T\, S_{lj}^{-1}\, \nu_{lj}}, \tag{6}$$

with ν_{lj} being the innovation and S_{lj} being the innovation covariance between a point $\mathbf{z}_l \in \mathcal{Z}_i$ and contour point $\mathbf{x}_{\mathcal{P}_j}$

of the predicted set $\mathcal{P}_j$ transformed into the sensor frame. More precisely, these two terms are given as

$$\nu_{lj} = \mathbf{z}_l - (H\mathbf{x}_j(k+1|k) + \mathbf{x}_{\mathcal{P}_j}) \tag{7}$$
$$S_{lj} = H\, C_j(k+1|k)\, H^T + R_l , \tag{8}$$

where $H = \left(\begin{smallmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{smallmatrix}\right)$ is the measurement Jacobian and R_l the $2{\times}2$ observation covariance whose entries reflect the noise in the measurement process of the range finder.

The probability that cluster $\mathcal{Z}_i$ originates from group track G_j is finally given by a zero-centered Gaussian,

$$\mathcal{N}_i = \frac{1}{2\pi\sqrt{\det(S_{lj})}} \exp\left(-\tfrac{1}{2} d^2_{\text{HD}}(\mathcal{Z}_i,\, G_j)\right) . \tag{9}$$

3.5 Group-to-Group Assignment Probability

To determine the probability that two groups G_i and G_j merge, we compute the distance between their closest contour points in a Mahalanobis sense. In doing so, we have to account for the clustering distance d_P, since we consider G_i and G_j to be one group as soon as their contours come closer than d_P. Let $\Delta\mathbf{x}_{\mathcal{P}_{ij}} = \mathbf{x}_{\mathcal{P}_i} - \mathbf{x}_{\mathcal{P}_j}$ be the vector difference of two contour points of G_i and G_j, respectively. We then subtract d_P from $\Delta\mathbf{x}_{\mathcal{P}_{ij}}$ unless $\Delta\mathbf{x}_{\mathcal{P}_{ij}} \leq d_P$ for which $\Delta\mathbf{x}_{\mathcal{P}_{ij}} = \mathbf{0}$. Concretely, the modified difference becomes $\Delta\mathbf{x}'_{\mathcal{P}_{ij}} = \max(\mathbf{0},\ \Delta\mathbf{x}_{\mathcal{P}_{ij}} - d_P\, \mathbf{u}_{\mathcal{P}_{ij}})$ where $\mathbf{u}_{\mathcal{P}_{ij}} = \Delta\mathbf{x}_{\mathcal{P}_{ij}} / |\Delta\mathbf{x}_{\mathcal{P}_{ij}}|$.

To obtain a similarity measure that accounts for nearness of group contours *and* similar velocity, we augment $\Delta\mathbf{x}'_{\mathcal{P}_{ij}}$ by the difference in the velocity components,

$$\Delta\mathbf{x}^*_{\mathcal{P}_{ij}} = (\Delta\mathbf{x}'^{T}_{\mathcal{P}_{ij}},\ \dot{x}_i - \dot{x}_j\ ,\ \dot{y}_i - \dot{y}_j)^T . \tag{10}$$

We now determine the statistical compatibility of two groups G_i and G_j according to the four-dimensional minimum Mahalanobis distance

$$d^2_{\text{min}}(G_i,\, G_j) = \min_{\substack{\mathbf{x}_{\mathcal{P}_i} \in \mathcal{P}_i \\ \mathbf{x}_{\mathcal{P}_j} \in \mathcal{P}_j}} \left\{ D^2(\Delta\mathbf{x}^*_{\mathcal{P}_{ij}},\, C_i + C_j) \right\} . \tag{11}$$

The probability that two groups actually belong together, is finally given by

$$\mathcal{N}_{ij} = \frac{1}{(2\pi)^2\sqrt{\det(C_i + C_j)}} \exp\left(-\tfrac{1}{2} d^2_{\text{min}}(G_i,\, G_j)\right) . \tag{12}$$

In this formulation, only the upper-left $4{\times}4$ blocks of C_i and C_j are used, which excludes the group size estimate and the corresponding uncertainties from data association. In future work, these could be included as well.

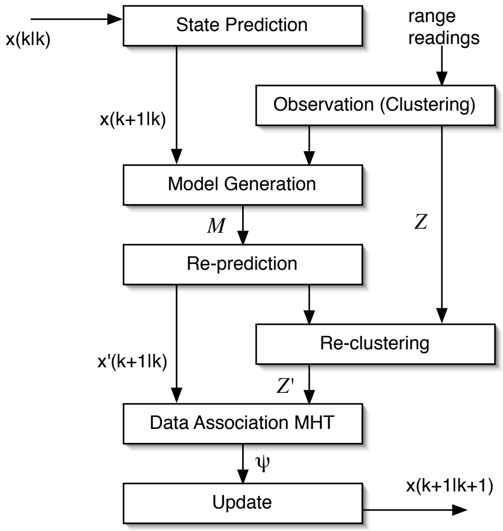

Fig. 3 Flow diagram of the tracking system. It differs from a regular tracker in the additional steps *model generation*, *re-prediction* and *re-clustering* (see explanations in section 4).

4 Tracking Cycle

This section describes the steps in the cycle of our Kalman filter-based group tracker. An overview is given by the flow diagram in Fig. 3. The structure differs from a regular tracker in the additional steps *model generation*, *track re-prediction*, and *re-clustering*.

- *State prediction:* In this step, the states of all existing group tracks are predicted under the assumption that they continue without interacting with other tracks, i.e., without splits or merges. See Sect. 3.3 for details.
- *Observation:* As described in Sect. 2, this step involves grouping the laser range data into clusters $\mathcal{Z}$.
- *Model Generation:* Models are generated based on the predicted group tracks and the clusters $\mathcal{Z}$, see Sect. 5.
- *Re-prediction:* Based on the model hypotheses that postulate a split, merge, or continuation event for each track, groups are re-predicted using these hypotheses so as to reflect the respective model, as explained in Sect. 3.3.
- *Re-clustering:* Re-clustering an observed cluster $\mathcal{Z}_i$ is necessary when it might have been produced by more than one group track, that is, it is in the gate of more than one track. If the model hypothesis postulates a merge for the involved tracks, nothing needs to be done. Otherwise, $\mathcal{Z}_i$ needs to be re-clustered, which is done using a nearest-neighbor rule: those points $\mathbf{z}_l \in \mathcal{Z}_i$ that share the same nearest neighbor track are combined to a new cluster. This step follows from the uniqueness assumption, which is common in target tracking and according to which an observation can only be produced by a single target.

- *Data Association MHT:* This step involves the generation, probability calculation, and pruning of data association hypotheses that assign re-predicted group tracks to re-clustered observations. See Sect. 6.
- *Update:* Each group track G_j that has been assigned to a cluster $\mathcal{Z}_i$ is updated with a standard linear Kalman filter. We use an observation vector $\tilde{\mathbf{z}}_i = (\bar{\mathbf{z}}_i, n_{\text{hs}}(\mathcal{Z}_i))^T$, that contains both the centroid position $\bar{\mathbf{z}}_i$ of $\mathcal{Z}_i$ and the number of human-sized blobs $n_{\text{hs}}(\mathcal{Z}_i)$ in the cluster. The update is then given by

$$\mathbf{x}(k+1|k+1) = \mathbf{x}(k+1|k) + K\Big(\tilde{\mathbf{z}}_i - \tilde{H}\mathbf{x}(k+1|k)\Big) \quad (13)$$

$$C(k+1|k+1) = C(k+1|k) - K\tilde{H}\, C(k+1|k) \quad (14)$$

with K being the Kalman gain matrix and $\tilde{H}$ the corresponding measurement Jacobian,

$$K = C(k+1|k) \cdot\ \tilde{H}^T\Big(\tilde{H}C(k+1|k)\tilde{H}^T + R_l\Big)^{-1} \quad (15)$$

$$\tilde{H} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \quad (16)$$

The contour points in $\mathcal{P}_j$ are replaced by the points in $\mathcal{Z}_i$ after being transformed into the reference frame of the posterior state $\mathbf{x}(k+1|k+1)$, as described in Sect. 3.2. Thereby, $\mathcal{P}_j$ always contains the most recent approximation of the group.

5 Model Generation and Model Probability

A model is defined to be a partitioning of tracks into groups. It assumes a particular state of the group formation process. New models, whose generation is described in this section, hypothesize about the evolution of that state. As this happens recursively, that is, based on the previous model of the last time index, the problem can thus be seen as a recursive clustering problem.

The space of possible model transitions is large since each group track can split into an unknown number of new tracks, or merge with an unknown number of other tracks. We therefore impose the gating condition for observations and tracks using the minimum average Hausdorff distance, thereby implementing a data-driven aspect into the model generation step:

- Multiple group tracks G_i can merge into one track only if there is an observation which is statistically compatible with all G_i.
- A group track can only split into multiple tracks that are all matched with observations in that very time step. Splits into occluded or obsolete tracks are not allowed.

Gating and statistical compatibility are both determined on a significance level α.

We further bound the possible number of model transitions as we assume that merge and split are binary operators. More precisely, we assume:

- At most two group tracks G_i, G_j can merge into one track at the same time.
- A track G_i can split at most into two tracks in one frame.
- A group track can not be involved in a split and a merge action at the same time.

These limitations are justified by the assumption that we observe the world much faster than the rate with which it evolves. This fact alleviates the impact of violations of the above assumptions: even if, for instance, a group splits into three subgroups at once, the tracker requires only two cycles to reflect this change.

A new model now defines for each group track if it is continued, split, or merged with another group track. The probability of a model is calculated using the constant prior probabilities for continuations and splits, p_C and p_S respectively, and the probability for a merge between two tracks G_i and G_j as $p_G \cdot \mathcal{N}_{ij}$. The latter term consists of a constant prior probability p_G and the group-to-group assignment probability $\mathcal{N}_{ij}$ defined in Sect. 3.5. Let N_C and N_S be the number of continued tracks and the number of split tracks in model M respectively, then the probability of M conditioned on the parent hypothesis Ω^{k-1} is

$$P(M|\Omega^{k-1}) = p_C^{N_C} \cdot p_S^{N_S} \prod_{G_i, G_j \in\, \Omega^{k-1}} (p_G \cdot \mathcal{N}_{ij})^{\delta_{ij}} \quad (17)$$

with δ_{ij} being 1 if G_i, G_j merge and 0 otherwise.

6 Multi-Model MHT

In this section we describe our adaptions and extensions of the original MHT by Reid [13] to a multi-model tracking approach that hypothesizes over both, data associations and models (as defined in the previous sections).

Let Ω_i^k be the i-th hypothesis at time k and $\Omega_{p(i)}^{k-1}$ its parent. Let further $\psi_i(k)$ denote a set of assignments that associate predicted tracks in $\Omega_{p(i)}^{k-1}$ to observations in $Z(k)$. As there are many possible assignment sets given $\Omega_{p(i)}^{k-1}$ and $Z(k)$, there are many children that can branch off a parent hypothesis, each with a different $\psi(k)$. This makes up an exponentially growing hypothesis tree.

The multi-model MHT introduces an intermediate tree level for each time step, on which models spring off from parent hypotheses (Fig. 4). In each model branch, the tracks of the parent hypothesis are first re-predicted to implement that particular model and then assigned to the (re-clustered)

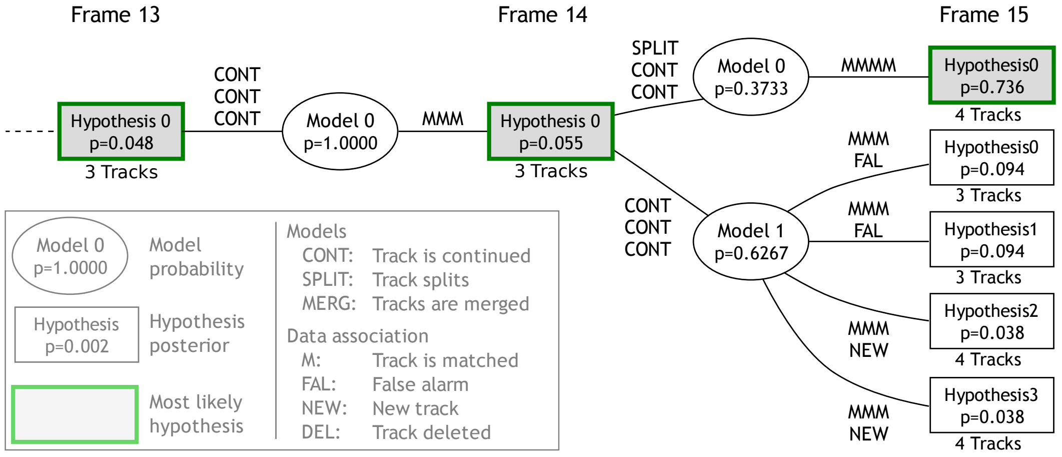

Fig. 4 The multi-model MHT. For each parent hypothesis, model hypotheses (ellipses) branch out and create their own assignment problems. In our application, models define which tracks of the parent hypothesis are continued, split, or merged. The tree shows frames 13 to 15 of figure 6. The split of group 1 between frames 14 and 15 is the most probable hypothesis after data association following model branch 0, although the continuation following model branch 1 is more probable (see the legend for details).

observations. Possible assignments for observations are existing tracks that *match* with existing tracks, *false alarms* or *new tracks*. Using the generalized formulation of Arras *et al.* [7] to deal with more than two track interpretation labels, tracks are interpreted as *matched*, *obsolete*, or *occluded*.

6.1 Assignment Set and Hypothesis Probability

The probability of a hypothesis in the multi-model MHT is calculated as follows. We compute the probability of a child hypothesis Ω_i^k given the observations from all time steps up to k, denoted by Z^k. According to the Markov assumption, it is the joint probability of the assignment set $\psi_i(k)$, the model M, and the parent hypothesis $\Omega_{p(i)}^{k-1}$, conditioned on the current observation $Z(k)$. Using Bayes rule, this can be expressed as the product of the data likelihood with the joint probability of assignment set, model and parent hypothesis

$$\begin{aligned}&P(\Omega_i^k|Z^k)\\&= P(\psi, M, \Omega_{p(i)}^{k-1}|Z(k)) \qquad (18)\\&= \eta \cdot P(Z(k)|\psi, M, \Omega_{p(i)}^{k-1}) \cdot P(\psi, M, \Omega_{p(i)}^{k-1}). \qquad (19)\end{aligned}$$

By using conditional probabilities, the third term on the right hand side can be factorized into the probabilities of the assignment set, the model, and the parent hypothesis

$$\begin{aligned}&P(\psi, M, \Omega_{p(i)}^{k-1})\\&= P(\psi|M, \Omega_{p(i)}^{k-1}) \cdot P(M|\Omega_{p(i)}^{k-1}) \cdot P(\Omega_{p(i)}^{k-1}). \qquad (20)\end{aligned}$$

The third factor in this product is known from the previous iteration, whereas the second factor represents the model probability derived in Sect. 5.

It remains to specify the first factor which is the probability of the assignment set ψ. The set ψ contains the assignments of observed clusters $\mathcal{Z}_i$ and group tracks G_j either to each other or to one of their possible labels listed above. Assuming independence between observations and tracks, the probability of the assignment set is the product of the individual assignment probabilities. Namely, they are p_M for matched tracks, p_F for false alarms, p_N for new tracks, p_O for tracks found to be occluded, and p_T for obsolete tracks scheduled for termination. If the number of new tracks and false alarms follow a Poisson distribution (as assumed by Reid [13]), the probabilities p_F and p_N have a sound physical interpretation as $p_F = \lambda_F V$ and $p_N = \lambda_N V$, where λ_F and λ_N are the average rates of events per volume multiplied by the observation volume V (the field of view of the sensor). The probability for an assignment ψ given a model M and a parent hypothesis Ω^{k-1} is then computed as

$$\begin{aligned}&P(\psi|M, \Omega^{k-1})\\&= p_M^{N_M}\, p_O^{N_O}\, p_T^{N_T}\, \lambda_F^{N_F} \lambda_N^{N_N}\, V^{N_F+N_N}, \qquad (21)\end{aligned}$$

where the Ns are the number of assignments to the respective labels in ψ.

Thanks to the independence assumption, also the data likelihood $P(Z(k)|\,\psi, M, \Omega_{p(i)}^{k-1})$ is computed by the product of the individual likelihoods of each observation cluster $\mathcal{Z}_i$ in $Z(k)$. If ψ assigns an observation $\mathcal{Z}_i$ to an existing track, we assume the likelihood of $\mathcal{Z}_i$ to follow a normal distribution, given by Eq. 9. Observations that are interpreted as false alarms and new tracks are assumed to be uniformly distributed over the observation volume V, yielding a likelihood of $1/V$. The data likelihood then becomes

$$P(Z(k)|\psi, M, \Omega^{k-1}) = \left(\tfrac{1}{V}\right)^{N_N+N_F} \prod_{i=1}^{N_{\mathcal{Z}}} \mathcal{N}_i^{\,\delta_i}, \qquad (22)$$

where δ_i is 1 if $\mathcal{Z}_i$ has been assigned to an existing track, and 0 otherwise.

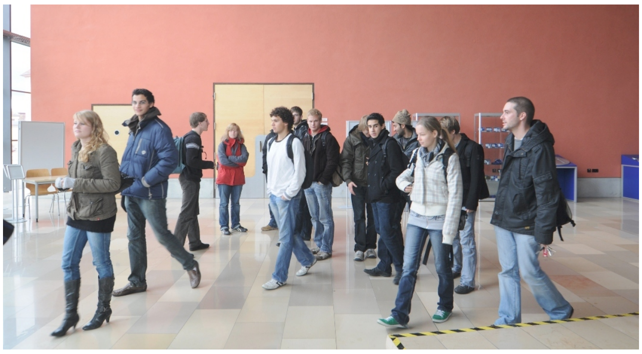

Fig. 5 Space where we have recorded the datasets for our experiments.

Substitution of Eqs. (17), (21), and (22) into Eq. (18) leads, like in the original MHT approach, to a compact expression, independent on the observation volume V.

Finally, normalization is performed yielding a true probability distribution over the child hypotheses of the current time step. This distribution is used to determine the current best hypothesis and to guide the pruning strategies.

6.2 Hypothesis Pruning

Pruning is essential in implementations of the MHT algorithm, as otherwise the number of hypotheses grows boundless. The following strategies are employed:

K-best branching: instead of creating all children of a parent hypothesis, the algorithm proposed by Murty [16] generates only the K most probable hypotheses in polynomial time. We use the multi-parent variant of Murty's algorithm, mentioned in [22], that generates the global K best hypotheses for all parents.

Ratio pruning: a lower limit on the ratio of the current and the best hypothesis is defined. Unlikely hypotheses with respect to the best one, being below this threshold, are deleted. Ratio pruning overrides K-best branching in the sense that if the lower limit is reached earlier, less than K hypotheses are generated.

N-scan back: the N-scan-back algorithm considers an ancestor hypothesis at time $k-N$ and looks ahead in time onto all children at the current time k (the leaf nodes). It only keeps the subtree at $k-N$ with the highest sum of leaf node probabilities. All other branches at $k-N$ are discarded.

More details on these pruning strategies can be found in the work of Cox and Hingorani [14].

7 Experiments

To analyze the performance of our system, we collected two data sets in the entrance hall of a university building, shown in Fig. 5. We used a Pioneer II robot equipped with a SICK laser scanner mounted at 30 cm above floor, scanning at 10 fps. In two unscripted experiments (dataset 1 with a stationary robot and dataset 2 with a moving robot), up to 20 people are in the field of view of the sensor. They form a large variety of groups during social interaction, move around, stand together and jointly enter and leave the hall, see Fig. 6.

Table 1 Summary of the two datasets used in the experiments.

	Dataset 1	Dataset 2
Number of frames	578	991
Avg. / max people	6.25 / 13	8.99 / 20
Avg. / max groups	2.60 / 4	4.16 / 8
Number of splits / merges	5 / 10	48 / 44
Number of new tracks / deletions	19 / 15	34 / 39

To obtain ground truth information, we labeled each single range reading. Beams that belong to a person receive a person-specific label, other beams are labeled as non-person. These labels are kept consistent over the entire duration of the data sets. People that socially interact with each other (derived by observation) are said to belong to a group with a group-specific label. Summed over all frames, the ground truth contains 5,629 labeled groups and 12,524 labeled people.[1] For further details, see Tab. 1.

The ground truth data is used for performance evaluation and to learn the parameters of our tracker. The values, determined by counting the related events in the ground truth and dividing by the total number of these events, are $p_M = 0.79$, $p_O = 0.19$, $p_T = 0.02$, $p_F = 0.06$, $p_N = 0.02$ for the data association probabilities, and $p_C = 0.63$, $p_S = 0.16$, $p_G = 0.21$ for the group formation probabilities. When evaluating the performance of the tracker, we separated the data into a training set and a validation set to avoid overfitting.

The state uncertainty for new tracks is given by $\sigma_x = \sigma_y = 0.1$, $\sigma_{\dot{x}} = \sigma_{\dot{y}} = 0.5$, and $\sigma_n = 0.2$. The noise parameter for the motion model are given by $\epsilon_x = \epsilon_y = 0.2$, $\epsilon_{\dot{x}} = \epsilon_{\dot{y}} = 0.3$, and $\epsilon_n = 0.1$.

Six frames of the current best hypothesis from the second dataset are shown in Fig. 6. The corresponding hypothesis tree for frame 15 is shown in Fig. 4. The sequence exemplifies movement and formation of several groups.

7.1 Clustering Error

This section analyzes how well the presented group tracker can recover the true group formation processes, i.e., which people actually belong together according to their social interaction as encoded in the ground truth.

We compute the clustering error of the tracker using the ground truth information on a per-beam basis. This is done

[1] Data sets, ground truth and result videos are available online at http://www.informatik.uni-freiburg.de/~lau/grouptracking

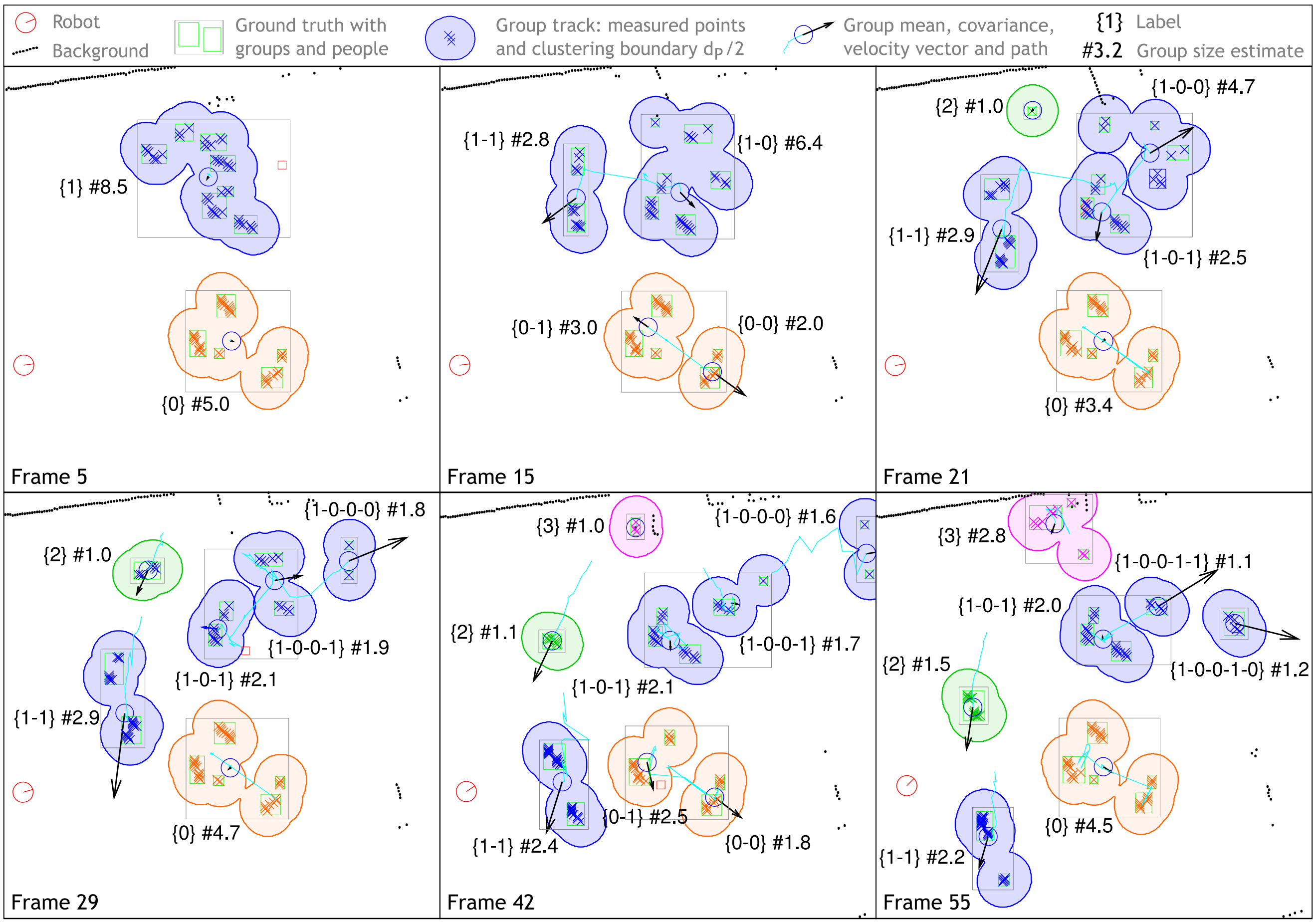

Fig. 6 Tracking results from the second data set. In frame 5, two groups are present. In frame 15, the tracker has correctly split group 1 into 1-0 and 1-1 (see Fig. 4). Between frames 15 and 29, group 1-0 has split up into groups 1-0-0 and 1-0-1 and split up again. New groups, labeled 2 and 3, enter the field of view in frames 21 and 42 respectively.

by counting how often set of points $\mathcal{P}$ of a track contains too many or wrong points (under-segmentation) and how often $\mathcal{P}$ is missing points (over-segmentation). Two examples for over-segmentation errors can be seen in Fig. 6, where group 0 and group 1-0 are temporarily over-segmented, compared to the ground truth which is visualized with a rectangle. However, from the history of group splits and merges stored in the group labels, the correct group relations can be determined in such cases.

For the first dataset, the clustering error rates for under-segmentation, over-segmentation, and the sum of both are shown in Fig. 7 (left), plotted against the clustering distance d_P.

We compare the clustering performance of our group tracker with a memory-less group clustering approach, which performs single-linkage clustering of the range data as described in Sect. 2 without using a tracking framework. The result is shown in Fig. 7 (middle).

The minimum clustering error of 3.1% is achieved by the tracker at $d_P = 1.3\,m$. The minimum error for the memory-less clustering is 7.0%, more than twice as high. In the second dataset, the error rates are higher due to the larger number of occlusions and the increased complexity in group interactions. Here, the minimum clustering error of the tracker is 9.6% while the error of the memory-less clustering reaches 20.2%, again more than twice as high.

To further investigate situations where tracking results differ from memory-less clustering, we recorded laser data of groups of people walking and passing in a corridor. An example is shown in Fig. 8, where one person passes between a group of two people. The memory-less approach would merge them immediately while the tracking approach, accounting for the velocity information, correctly keeps the groups apart by using re-clustering. This result shows that the group tracking problem is a *recursive* clustering problem that requires integration of information over time.

In the light of the proxemics theory the result of a minimal clustering error at 1.3 m is noteworthy. The theory predicts that when people interact with friends, they maintain a range of distances between 45 to 120 cm called personal space. When engaged in interaction with strangers, this distance is larger. As our data contains students who tend to

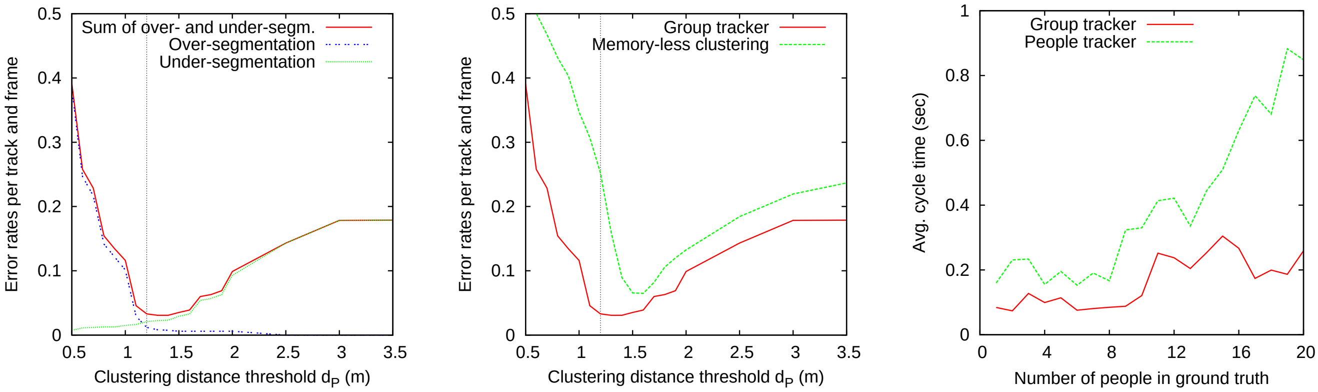

Fig. 7 *Left:* Clustering error of the group tracker as the sum of over-segmentation and under-segmentation error. The smallest error is achieved for a cluster distance of 1.3 m which is very close to the border of personal and social space according to the proxemics theory, marked at 1.2 m by the vertical line. *Middle:* Clustering error of the group tracker compared to memory-less single linkage clustering (without tracking). *Right:* Average cycle time for the group tracker versus a tracker for individual people plotted against the ground truth number of people.

know each other well, the result appears consistent with the findings of Hall.

7.2 Tracking Efficiency

When tracking groups of people rather than individuals, the assignment problems in the data association stage are of course smaller. At the same time, the introduction of an additional tree level, on which different models hypothesize over different group formation processes, comes with additional computational costs. We therefore compare our system with a person-only tracker realized by inhibiting all split and merge operations and reducing the cluster distance d_P to the value that yields the lowest error for clustering single people given the ground truth. For the second dataset, the resulting average cycle times versus the ground truth number of people is shown in Fig. 7 (right). The plots are averaged over different k from the range of 2 to 200 at a scan-back depth of $N = 30$.

With an increasing number of people, the cycle time for the people tracker grows much faster than the cycle time of the group tracker. Interestingly, even for small numbers of people the group tracker is faster than the people tracker. This is due to occasional over-segmentation of people into individual legs tracks. Also, as mutual occlusion of people in densely populated environments occurs frequently, the people tracker has to maintain many more occluded tracks than the group tracker, as occlusion of entire groups is rare. Also, the additional complexity of multiple models in the group tracker virtually disappears when the tracks are isolated due to the data-driven model generation.

This result clearly shows that our group tracking approach is more efficient. With an average cycle time of around 100 ms for up to 10 people on a Pentium IV at 3.2 GHz, the algorithm runs in real-time even with a non-optimized implementation.

7.3 Group Size Estimation

To evaluate the accuracy of our group size estimation approach, we define the error as the absolute difference between the estimated number of people in a group and the true value according to the labeled ground truth. For counting the number of human-sized clusters in a group as described in Sect. 2, a clustering threshold $d_{\text{hs}} = 0.3\,m$ is used.

For the first dataset, we find that the average error in group size estimation is 0.23 people with a standard deviation of 0.30. In the more complex dataset 2, the average error is 0.33 people with a standard deviation of 0.49. If the estimated group sizes are rounded to integers, the tracker determined the correct value in 88.9% of all cases in dataset 1 and in 84.3% for dataset 2.

If only deviations of more than one person are considered an error, the system was correct in 99.5% of all cases in dataset 1 and 97.5% in dataset 2.

8 Conclusion

In this paper, we presented a multi-model hypothesis tracking approach to track groups of people. We extended the original MHT approach to incorporate model hypotheses that describe track interaction events that go beyond what data association can express. In our application, models encode the formation of groups during split, merge, and continuation events. We further introduced a representation of groups that includes their shape, and employed the minimum average Hausdorff distance to account for the shape information when calculating association probabilities.

The proposed tracker has been implemented and tested using a mobile robot equipped with a laser range finder. The experiments with up to 20 people forming groups of different sizes demonstrated that the system is able to robustly track groups of people as they undergo complex for-

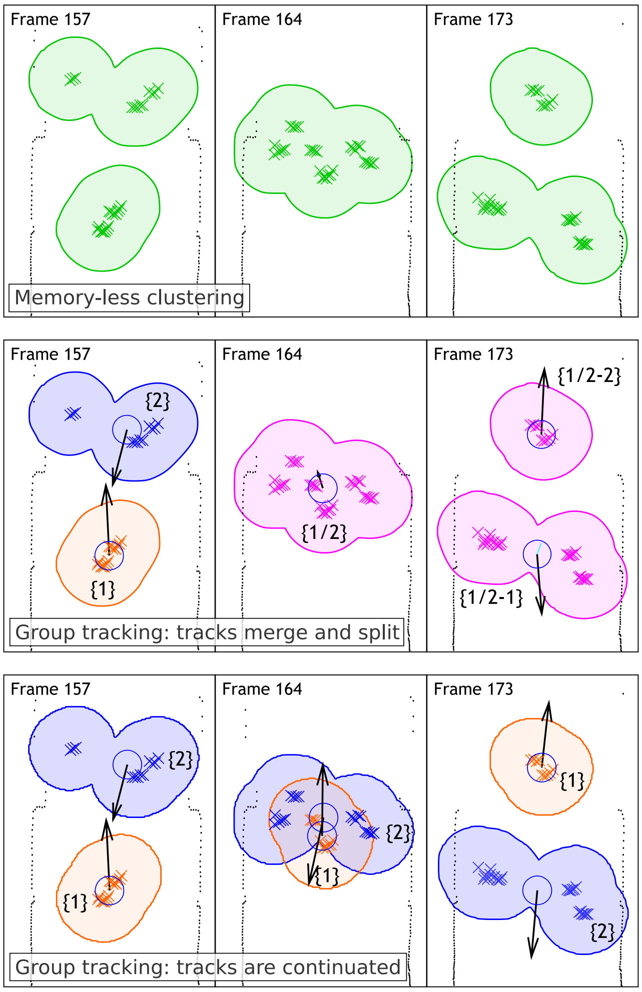

Fig. 8 One person crosses a group of two people. Since the groups interweave, memory-less clustering (top) unifies the two groups. Our group tracker can also create a model that postulates a merge of the groups, followed later by a split (middle). However, the model hypothesis leading to the most probable hypothesis in this situation continues both tracks and triggers re-clustering (see Sect. 4). This way, the crossing groups are tracked correctly (bottom). For a legend, see Fig. 6.

mation processes. Given ground truth data reflecting true interactions of people with over 12,000 labeled occurrences of people and groups, the experiments showed that the tracker could reproduce such processes with a low clustering error and estimate the number of people in groups with high accuracy. They also showed that in comparison with a memory-less single-frame clustering, our system performs significantly better in determining which people form a group.

The experiments demonstrated the ability of the approach to recover the actual social grouping of interacting people when compared to the ground truth. It was further found that the clustering threshold for detection that produces the best tracking results appears consistent with the proxemics theory. Finally, we showed that tracking groups of people is clearly more efficient than tracking individual people.

On a larger scale, we believe that group tracking is essential for robots to reason about social relations of people for various tasks in social robotics or human-robot interaction.

Acknowledgements This work was partly funded by the European Commission under contract number FP6-IST-045388 (INDIGO), by the German Research Foundation (DFG) under contract number SFB/TR-8, and by the German Federal Ministry of Education and Research (BMBF) under grant number 01IME01D (DESIRE).

References

1. B. Kluge, C. Köhler, and E. Prassler, "Fast and robust tracking of multiple moving objects with a laser range finder," in *Proceedings of the IEEE Int. Conf. on Robotics and Automation*, 2001.
2. A. Fod, A. Howard, and M. J. Mataric, "Laser-based people tracking," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, Washington DC, May 2002, pp. 3024–3029.
3. D. Schulz, W. Burgard, D. Fox, and A. Cremers, "People tracking with a mobile robot using sample-based joint probabilistic data association filters," *Intl. J. of Robotics Research (IJRR)*, vol. 22, no. 2, 2003.
4. J. Cui, H. Zha, H. Zhao, and R. Shibasaki, "Tracking multiple people using laser and vision," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Alberta, Canada, 2005.
5. W. Zajdel, Z. Zivkovic, and B. Kröse, "Keeping track of humans: Have I seen this person before?" in *IEEE International Conference on Robotics and Automation*, Barcelona, Spain, 2005.
6. G. Taylor and L. Kleeman, "A multiple hypothesis walking person tracker with switched dynamic model," in *Proc. of the Australasian Conference on Robotics and Automation*, Canberra, Australia, 2004.
7. K. O. Arras, S. Grzonka, M. Luber, and W. Burgard, "Efficient people tracking in laser range data using a multi-hypothesis leg-tracker with adaptive occlusion probabilities," in *IEEE International Conference on Robotics and Automation (ICRA)*, Pasadena, CA, USA, May 2008.
8. Z. Khan, T. Balch, and F. Dellaert, "MCMC data association and sparse factorization updating for real time multitarget tracking with merged and multiple measurements," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, 2006.
9. M. Mucientes and W. Burgard, "Multiple hypothesis tracking of clusters of people," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, October 2006, pp. 692–697.
10. S. McKenna, S. Jabri, Z. Duric, A. Rosenfeld, and H. Wechsler, "Tracking groups of people," *Computer Vision and Image Understanding*, vol. 80, no. 1, pp. 42–56, October 2000.
11. G. Gennari and G. D. Hager, "Probabilistic data association methods in visual tracking of groups," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2004.
12. B. Bose, X. Wang, and E. Grimson, "Multi-class object tracking algorithm that handles fragmentation and grouping," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2007, pp. 1–8.
13. D. B. Reid, "An algorithm for tracking multiple targets," *IEEE Transactions on Automatic Control*, vol. AC-24, no. 6, pp. 843–854, 1979.
14. I. Cox and S. Hingorani, "An efficient implementation of Reid's multiple hypothesis tracking algorithm and its evaluation for the purpose of visual tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 2, pp. 138–150, 1996.

15. S.-W. Joo and R. Chellappa, “A multiple-hypothesis approach for multiobject visual tracking,” *IEEE Transactions on Image Processing*, vol. 16, no. 11, pp. 2849–2854, November 2007.
16. K. Murty, “An algorithm for ranking all the assignments in order of increasing cost,” *Operations Research*, vol. 16, pp. 682–687, 1968.
17. B. Lau, K. O. Arras, and W. Burgard, “Tracking groups of people with a multi-model hypothesis tracker,” in *International Conference on Robotics and Automation (ICRA)*, Kobe, Japan, 2009.
18. E. Hall, *Handbook of Proxemics Research*. Society for the Anthropology of Visual Communications, 1974.
19. K. O. Arras, Óscar Martínez Mozos, and W. Burgard, “Using boosted features for the detection of people in 2d range data,” in *Proc. IEEE Intl. Conf. on Robotics and Automation (ICRA’07)*, Rome, Italy, 2007.
20. J. Hartigan, *Clustering Algorithms*. John Wiley & Sons, 1975.
21. M. P. Dubuisson and A. K. Jain, “A modified Hausdorff distance for object matching,” in *Intl. Conference on Pattern Recognition*, vol. 1, Jerusalem, Israel, 1994, pp. A:566–568.
22. I. Cox and M. Miller, “On finding ranked assignments with application to multi-target tracking and motion correspondence,” *IEEE Trans. on Aerospace and Electronic Systems*, vol. 31, no. 1, pp. 486–489, 1995.

Boris Lau is a research scientist in the Autonomous Intelligent Systems Group at the University of Freiburg (Germany). His research interests are people detection and tracking and robot navigation in populated environments. He received his diploma degree in computer science from the Technical University of Ilmenau in 2007. His prior work comprises vision based robotics at the Laboratory for Active and Attentive Vision at York University (Canada), as well as visual tracking and developmental modeling at the Cognitive Systems and Cognition Laboratory at University of California, San Diego (USA).

Kai Oliver Arras received his Masters in EE from ETH Zurich in 1996 and his Dr. degree from EPFL in 2003. After a post-doctoral stay at the Center for Autonomous Systems, KTH Stockholm, he founded Nurobot Automation and Artefacts, mainly active in industrial navigation systems. In 2006 he joined the Autonomous Intelligent Systems Group at the University of Freiburg as a post-doc, and in 2007, he joined Evolution Robotics, Pasadena, as a Senior Research Scientist. Since July 2008 he holds a Junior Research Group Leader position at the University of Freiburg and is head of the Social Robotics Laboratory in the Department of Computer Science.

Wolfram Burgard studied computer science at the University of Dortmund, Germany, and received is Ph.D. degree from the Department of Computer Science of the University of Bonn in 1991. In 1999 Wolfram Burgard became professor at the Department of Computer Science of the University of Freiburg where he heads the research laboratory for Autonomous Intelligent Systems. His areas of interest lie in artificial intelligence and robotics. They cover various aspects of mobile robotics including mobile robot navigation, multi-robot systems, state estimation, robot learning, and human robot interaction.