Efficient People Tracking in Laser Range Data using a Multi-Hypothesis Leg-Tracker with Adaptive Occlusion Probabilities

Kai O. Arras Slawomir Grzonka Matthias Luber Wolfram Burgard

***Abstract*— We present an approach to laser-based people tracking using a multi-hypothesis tracker that detects and tracks legs separately with Kalman filters, constant velocity motion models, and a multi-hypothesis data association strategy. People are defined as high-level tracks consisting of two legs that are found with little model knowledge. We extend the data association so that it explicitly handles track occlusions in addition to detections and deletions. Additionally, we adapt the corresponding probabilities in a situation-dependent fashion so as to reflect the fact that legs frequently occlude each other. Experimental results carried out with a mobile robot illustrate that our approach can robustly and efficiently track multiple people even in situations of high levels of occlusion.**

I. INTRODUCTION

People tracking is a key technology for robots that operate in populated environments. Knowledge about presence, position, and motion state of people will enable robots to better understand and anticipate their intentions and actions. Apart from human-robot interaction and cooperation scenarios, applications of laser-based people tracking include also surveillance, crowd control, or pedestrian detection for intelligent cars.

In this paper we consider the problem of people tracking from data acquired with two-dimensional laser range finders. In most related work on laser-based people tracking [1], [2], [3], [4], [5], [6], [7], [8], [9], a person is represented as a single state that encodes torso position and velocities. People are extracted from range data as single blobs or found by merging nearby point clusters that correspond to legs. However, the appearance of people in laser range data depends on the mounting height of the sensor: at hip height a human torso is typically seen as a single local-minimum blob while at foot height, legs produce separate, fast-moving smaller blobs. In practice, the mounting height of the sensor is often constrained by the application or the robot's form factor and not only by the researcher to suit the needs of a tracking algorithm. Safety regulations, for instance, require laser scanners to be mounted at foot height. Obviously, at this height, modeling people as single blobs can be problematic and thus motivates leg tracking as an approach to laser-based people tracking. Accordingly, the problem of people tracking has recently been addressed as a leg tracking problem [10], [11] where people are represented by the states of two legs, either in a single augmented state [11] or as a high-level track to which two low-level leg tracks are associated [10].

All authors are with the University of Freiburg, Department of Computer Science, Georges-Koehler-Allee 79, 79110 Freiburg, Germany {arras,grzonka,luber,burgard}@informatik.uni-freiburg.de

Multi-hypothesis tracking (MHT) [12], [13] 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, confirmation, occlusion, and deletion. Other multi-target data association techniques such as the nearest neighbor filter, the track splitting filter, or the JPDAF are less powerful or sub-optimal in nature [14].

In the context of people tracking with laser range finders, Taylor *et al.* [10] employ an MHT to resolve ambiguities in the problem of fitting a walking person into two leg measurements. The authors use a geometric occlusion model to decrease the detection probability if an occlusion is to be expected. Mucientes *et al.* [9] cluster people into groups and utilize an MHT to handle the assignments of measurements to single tracks and clusters. Given the high-level concept of groups, additional assignments of measurements to tracks within groups become possible for which the authors derive appropriate probabilities.

In this paper we track legs of people and utilize a multiple hypothesis tracking approach for data association. Opposed to most related work in the laser-based people tracking literature, we address the problem of tracking legs that are measured individually. Based on the resulting leg tracks, we create person tracks using the multivariate weighted mean if two tracks are sufficiently close and move in the same direction for a certain time frame. Once a person track has been validated over time, we adapt the individual occlusion probabilities of both associated leg tracks to account for the fact that legs frequently occlude each other. To this end, we extend the MHT framework to explicitely take into account potential occlusions by introducing adaptive conditional assignment probabilities.

The paper is structured as follows. The next section briefly describes the Kalman filter-based tracker used for detecting and tracking legs. Section III reviews the multi-hypothesis tracking approach, especially the expressions to calculate the hypothesis probabilities. Section IV introduces the concept of person tracks and how they are found. This section also contains the derivation of the probability equations that are needed to adapt the occlusion probabilities of individual tracks. Section V describes the experimental results.

II. KF-BASED LEG TRACKER

This section describes the KF-based multi-target tracker that is used to track legs of people. We briefly go through the tracking cycle. For the details of Kalman filtering and target tracking the reader is referred to Bar-Shalom and Li [14].

State prediction. A leg track is represented as $\mathbf{x} = (x, y, v_x, v_y)$ where x and y are the track position and v_x and v_y the x and y components of the track velocity. With this state representation new tracks can be properly initialized with $v_x = v_y = 0$. For motion prediction, a constant velocity model is employed.

Measurement prediction. As the x- and y-coordinates of a track are directly observable, the 2×4 measurement matrix H is formed by the 2×2 identity matrix in x and y and the 2×2 zero matrix in v_x and v_y.

Observation. The observation step consists in detecting people in range data. The problem can be seen as a classification problem that consists in finding those laser beams that correspond to people and to discard other beams. Typically, hand designed classifiers have been employed for this task with a manual selection of features and thresholds. In a recent work we used AdaBoost, a supervised learning technique, to learn a classifier for groups of adjacent beams that correspond to people [15]. AdaBoost takes a labeled training set and a (possibly large) vocabulary of features that may or may not be appropriate for the given classification task. The method then creates a classifier by selecting the most informative features and finding the best thresholds (based on the training set). The AdaBoost classifier proved to be superior to a manually designed classifier. It shall therefore be used also in this work.

The observation step delivers the set of observations (or measurements) $\mathbf{z}_k = \{z_k^1, z_k^2, \ldots, z_k^{M_k}\}$ at time index k. M_k denotes the current number of measurements.

Data association. For data association we employ a modified MHT approach described in the sections hereafter.

Estimation. Given that both, the state and measurement prediction models are linear, a (non-extended) Kalman filter as the optimal estimator under the Gaussian assumption can be employed.

III. Multi Hypothesis Tracker

In this section we review the MHT as described in the two papers by Reid [12] and Cox *et al.* [13]. In the original paper by Reid [12] measurements can be interpreted as matches with existing tracks, new tracks, or false alarms. Tracks are interpreted as *detected* (when they match with a measurement) or *not detected*. Deletion of tracks is not handled by the MHT but by a heuristics based on sequences of consecutive non-detections. Cox *et al.* [13] extend this framework with the interpretation of tracks as *deleted*. Thereby, the MHT handles the entire life-cycle of tracks from creation and confirmation (by matching) to deletion and occlusion (which is non-detection and non-deletion).

In order to adapt the occlusion probabilities of individual leg tracks later in this paper, it is necessary to reconsider the derivation of the hypothesis probabilities in the MHT, especially the assignment set probabilities.

Let Ω_j^k be the j-th hypothesis at time k and $\Omega_{p(j)}^{k-1}$ the parent hypothesis from which Ω_j^k was derived. Let further $\Psi_j(k)$ denote a set of assignments that, based on the parent hypothesis $\Omega_{p(j)}^{k-1}$ and the current measurement $\mathbf{z}_k$, gives rise to Ω_j^k.

	$\mathbf{x}_1$	$\mathbf{x}_2$	$\mathbf{x}_{nt}$	$\mathbf{x}_{fa}$
z_1	0	0	1	0
z_2	1	0	0	0
z_{del}	0	1	0	0

TABLE I

Example of an assignment.

The assignment set $\Psi_j(k)$ associates each measurement either to an existing track, a false alarm, or a new track and marks a track as *detected* or *deleted*. Assignment sets are best visualized in matrix form such as the example in Table I that shows a set of assignments of tracks $\mathbf{x}_1, \mathbf{x}_2$ with measurements z_1 and z_2. An assignment is denoted by a non-zero entry in the matrix. The example shows a situation in which track $\mathbf{x}_1$ is assigned to measurement z_2, track $\mathbf{x}_2$ is scheduled for deletion, and measurement z_1 is interpreted as a new track.

There are as many possible assignment sets $\Psi_j(k)$ as we can distribute 1's and 0's over such matrices under the constraints of unique measurement-to-track associations and that the only zero-valued rows and columns can belong to the events *deletion*, *new track*, and *false alarm*. An assignment set has a probability that is determined by the probabilities of these events and the probability of a specific distribution of 1's and 0's.

Given an assignment set probability and the probability of the parent hypothesis $\Omega_{p(j)}^{k-1}$, we can calculate the probability of each child hypothesis that has been created as $\Psi_j(k)$. This calculation is done recursively [12]:

$$\begin{aligned} p(\Omega_j^k|\mathbf{z}_k) &= p(\Psi_j(k), \Omega_{p(j)}^{k-1}|\mathbf{z}_k) \\ &\overset{Bayes+}{\underset{Markov}{=}} \eta\, p(\mathbf{z}_k|\Psi_j(k), \Omega_{p(j)}^{k-1}) p(\Psi_j(k)|\Omega_{p(j)}^{k-1}) \cdot \\ &\quad\; p(\Omega_{p(j)}^{k-1}). \end{aligned} \tag{1}$$

The rightmost term on the right-hand side is the recursive term, that is, the probability of its parent. Factor η is a normalizer. The leftmost term on the right-hand side after the normalizer η is the measurement likelihood. We assume that a measurement z_k^i associated to a track $\mathbf{x}_j$ has a Gaussian pdf centered around the measurement prediction $\hat{z}_k^j$ with innovation covariance matrix $\mathcal{S}_k^{i,j}$, $\mathcal{N}(z_k^i) := \mathcal{N}(z_k^i\,; \hat{z}_k^j, \mathcal{S}_k^{i,j})$. We further assume the pdf of a measurement belonging to a new track or false alarm being uniform in the observation volume V (the field of view of the sensor) with probability V^{-1}. Thus

$$\begin{aligned} p(\mathbf{z}_k|\Psi_j(k), \Omega_{p(j)}^{k-1}) &= \prod_{i=1}^{M_k} \mathcal{N}(z_k^i)^{\delta_i} V^{1-\delta_i} \\ &= V^{-(N_{fal}+N_{new})} \prod_{i=1}^{M_k} \mathcal{N}(z_k^i)^{\delta_i} \end{aligned} \tag{2}$$

with N_{fal} and N_{new} the number of measurements labeled as *false alarms* and *new tracks* respectively. δ_i is an indicator variable being 1 if and only if measurement i has been associated to a track, 0 otherwise.

The central term on the right-hand side of Equation (1) is the probability of an assignment set, $p(\Psi_j(k)|\Omega^{k-1}_{p(j)})$, which is composed of three terms:

1) The probability of the *number* of tracks N_{det}, N_{fal}, N_{new} with a certain label. In Reid's case, with tracks being either labeled *detected* or *not detected*, the number of detected tracks N_{det} given the total number of tracks in the parent hypothesis, N, follows a binomial distribution

$$p(N_{det}|\Omega^{k-1}_{p(j)}) = \binom{N}{N_{det}} p_{det}^{N_{det}}(1-p_{det})^{(N-N_{det})} \tag{3}$$

Assuming that the *number* of false alarm and the *number* of new tracks both follow a Poisson distribution with expected number of events $\lambda_{fal}V$ and $\lambda_{new}V$ in the observation volume V respectively, we obtain

$$\begin{aligned} p(N_{det}, N_{fal}, N_{new}|\Omega^{k-1}_{p(j)}) = \\ \binom{N}{N_{det}} p_{det}^{N_{det}}(1-p_{det})^{(N-N_{det})} \\ \cdot\mu(N_{new};\lambda_{new}V)\cdot\mu(N_{fal};\lambda_{fal}V) \end{aligned} \tag{4}$$

where $\mu(n;\lambda V)$ is the Poisson distribution for n events when the average rate of events is λV.

2) The probability of a specific assignment of measurements so that $M_k = N_{det} + N_{fal} + N_{new}$ holds. The probability is determined as 1 over the number of combinations which is

$$\binom{M_k}{N_{det}}\binom{M_k-N_{det}}{N_{fal}}\binom{M_k-N_{det}-N_{fal}}{N_{new}} \tag{5}$$

where the last term equals 1.

3) The probability of a specific assignment of tracks given that a track can either be *detected* or *not detected*. The probability is determined as 1 over the number of these assignments

$$\frac{N!}{(N-N_{det})!}\binom{N-N_{det}}{N_{det}}. \tag{6}$$

The first term follows from the combinatorial fact, that a track can be chosen only once and the track-to-measurement order matters.

It is noteworthy (and one of the key contributions of Reid [12]) that in the product of these three probabilities many terms cancel out, and substituted into the Equation (1), the final probability $p(\Omega^k_j|\mathbf{z}_k)$ becomes a simple and easy to calculate expression independent of the observation volume V.

IV. Person Tracking and Occlusion Adaptation

The tracking system presented in the previous sections maintains N tracks that correspond to human legs. Only on the level of these N tracks, we reason on the existence of people by the use of the following model knowledge:

1) People have always two legs
2) Legs are close to each other
3) Legs move in a similar direction
4) Legs have a higher probability of occluding each other than being occluded by other people's legs or objects

In contrast to previous work [10], [11] we do not describe people by a more complex model that also encodes the dynamics of a walking person. People have a large variety of leg motion patterns (such as random steps on the spot while they are waiting) that are not adequately captured by walking models typically found in the literature [10].

To create a person track, we implement the above-mentioned model as follows:

1) A person track is defined as a high-level track to which two legs tracks are associated. The state of a person is estimated from the state of the two legs tracks using the multivariate weighted mean.
2) Two tracks $\mathbf{x}_i, \mathbf{x}_i$ that satisfy a nearness condition given a threshold θ_d which in our case is set to 0.75 meter form a person candidate.
3) A person candidate is validated if the two tracks maximize the scalar product of their orientations summed over the track histories $S = \sum_t < \theta^t_i, \theta^t_j >$ with $\theta_i = \mathrm{atan2}(v^2_{y,i}, v^2_{x,i})$ being the orientation of track $\mathbf{x}_i$. In practice, we calculate S only in a sliding window over the last L steps and validate a person track that satisfies $S > \theta_a$ where θ_a is an experimentally determined threshold.
4) The adaptation of the occlusion probability is described in detail in the following subsection.

Person tracks are deleted if either the MHT deletes one or both of its leg tracks or if condition 2) does not hold anymore for L consecutive steps.

A. Adaptation of occlusion probability

According to Reid [12], who only considers the label *detected*, the number of tracks with this label, N_{det}, follows a binomial distribution. In the more general case, in which we have an arbitrary number of labels, the number of tracks with a given label follows a *multinomial distribution*.

Besides *detection* (according to Reid [12]) and *deletion* (introduced by Cox and Hingorani [13]) we introduce the label *occlusion*. Thus, the pdf of the labeling of the tracks into *detected*, *occluded*, and *deleted* is

$$\begin{aligned} p(N_{det}, N_{occ}, N_{del}|\Omega^{k-1}_{p(j)}) = \\ \frac{N!}{N_{det}!N_{occ}!N_{del}!} p_{det}^{N_{det}} p_{occ}^{N_{occ}} p_{del}^{N_{del}} \end{aligned} \tag{7}$$

with $p_{det} + p_{occ} + p_{del} = 1$ and $N = N_{det} + N_{occ} + N_{del}$. Equation (7) is the generalization of Equation (3) and allows to specifically adjust the label probabilities. Occlusions are no longer implied by non-detection and non-deletion but are made explicit as a label with their own specific probability.

However, adjusting individual probabilities raises the question whether probabilities of assignments and hypotheses remain properly normalized across branches in the hypothesis tree. We will now verify that the consistency in this sense is maintained.

In our case, there are leg tracks that are associated to validated person tracks and leg tracks that are either associated to non-validated person tracks or to no person track at all. We will denote the former as *approved* (by the superscript A) and the latter as *free* (by the superscript F). With N^A and N^F as the number of approved and the number of free tracks respectively, $N = N^A + N^F$ and likewise

$$N^F = N^F_{det} + N^F_{occ} + N^F_{del} \tag{8}$$
$$N^A = N^A_{det} + N^A_{occ} + N^A_{del}. \tag{9}$$

The evidence *approved* and *free* conditions the probabilities in Equation (7) such that the right-hand side must be rewritten as the product of two multinomial distributions, each with three conditional probabilities $p_{det|F}, p_{del|F}, p_{occ|F}$ and $p_{det|A}, p_{del|A}, p_{occ|A}$ for which $p_{det|F} + p_{del|F} + p_{occ|F} = 1$ and $p_{det|A} + p_{del|A} + p_{occ|A} = 1$ must hold. The product of multinomial distributions is explained by the fact that a track can only be either approved or free.

As a consequence, the three product terms that compose the assignment set probability, $p(\Psi_j(k)|\Omega^{k-1}_{p(j)})$, are altered as follows. The first term, the probability of the *number* of tracks with a certain label becomes

$$\begin{aligned} p(N^F_{det}, N^F_{occ}, N^F_{del}, N^A_{det}, N^A_{occ}, N^A_{del}, N_{new}, N_{fal}|\Omega^{k-1}_{p(j)}) \\ = \frac{N^F!}{N^F_{det}!N^F_{occ}!N^F_{del}!} \cdot p^{N^F_{det}}_{det|F} \cdot p^{N^F_{occ}}_{occ|F} \cdot p^{N^F_{del}}_{del|F} \cdot \\ \frac{N^A!}{N^A_{det}!N^A_{occ}!N^A_{del}!} \cdot p^{N^A_{det}}_{det|A} \cdot p^{N^A_{occ}}_{occ|A} \cdot p^{N^A_{del}}_{del|A} \cdot \\ \mu(N_{fal}; \lambda_{fal}V) \cdot \mu(N_{new}; \lambda_{new}V) \end{aligned} \tag{10}$$

The second term, the probability of a specific combination of these numbers, is calculated as 1 over the number of these combinations, which is

$$\begin{aligned} &\binom{M_k}{N^F_{det}} \binom{M_k - N^F_{det}}{N^A_{det}} \binom{M_k - N^F_{det} - N^A_{det}}{N_{new}} \cdot \\ &\binom{M_k - N^F_{det} - N^A_{det} - N_{new}}{N_{fal}} \\ &= \frac{M_k!}{N^F_{det}!N^A_{det}!N_{new}!N_{fal}!} \end{aligned} \tag{11}$$

since $M_k = N^F_{det} + N^A_{det} + N_{new} + N_{fal}$.

Similarly, for the third term, the probability of the number of track-to-measurement associations determined as 1 over the number of these associations, is 1 over

$$\begin{aligned} &\frac{N^F!}{(N^F - N^F_{det})!} \binom{N^F - N^F_{det}}{N^F_{occ}} \binom{N^F - N^F_{det} - N^F_{occ}}{N^F_{del}} \cdot \\ &\frac{N^A!}{(N^A - N^A_{det})!} \binom{N^A - N^A_{det}}{N^A_{occ}} \binom{N^A - N^A_{det} - N^A_{occ}}{N^A_{del}} \\ &= \frac{N^F!N^A!}{N^F_{occ}!N^F_{del!}N^A_{occ}!N^A_{del}!} \end{aligned} \tag{12}$$

When combining these results, many terms cancel out like in Reid's approach [12]. Accordingly, we obtain the assignment set probability as

$$\begin{aligned} &p(\Psi_j(k)|\Omega^{k-1}_{p(j)}) = \\ &\eta' \cdot p^{N^F_{det}}_{det|F} \cdot p^{N^F_{occ}}_{occ|F} \cdot p^{N^F_{del}}_{del|F} \cdot p^{N^A_{det}}_{det|A} \cdot p^{N^A_{occ}}_{occ|A} \cdot p^{N^A_{del}}_{del|A} \cdot \\ &\lambda^{N_{new}}_{new} \cdot \lambda^{N_{fal}}_{fal} \cdot V^{N_{new}+N_{fal}} \end{aligned} \tag{13}$$

where η' is a constant normalization factor.

Substituting Equation (13) and the measurement likelihood from Equation (2) into Equation (1) yields the final expression for the probability of a child hypothesis

$$\begin{aligned} p(\Omega^k_j|\mathbf{z}_k) = \; &\eta'' \prod_{i=1}^{M_k} \mathcal{N}(z^i_k)^{\delta_i} \cdot \\ &p^{N^F_{det}}_{det|F} \cdot p^{N^F_{occ}}_{occ|F} \cdot p^{N^F_{del}}_{del|F} \cdot p^{N^A_{det}}_{det|A} \cdot p^{N^A_{occ}}_{occ|A} \cdot p^{N^A_{del}}_{del|A} \cdot \\ &\lambda^{N_{new}}_{new} \cdot \lambda^{N_{fal}}_{fal} \cdot p(\Omega^{k-1}_p(j)). \end{aligned} \tag{14}$$

Here $\eta'' = \eta \cdot \eta'$ is a constant normalization factor which ensures that the probabilities of the hypotheses Ω^k_j sum up to 1. It can be shown that η'' only depends on M_k. This means that within the same generation of hypotheses – for which M_k is identical – proper normalization across all branches in the tree, that is across all hypothesis probabilities, is guaranteed.

B. Branching and Pruning Strategies

For an efficient implementation of an MHT, pruning strategies that limit the exponential explosion of hypotheses are mandatory. As proposed by Cox and Hingorani [13] we make use of the following strategies:

- *k-Best Branching*. Instead of creating all children, we generate only the k best children for each parent hypothesis. This can be done in polynomial time with an algorithm proposed by Murty [16].
- *Ratio Pruning*. A lower limit on the ratio of the current and the best hypothesis is defined. Unlikely hypotheses being below this threshold are deleted.
- *N-scan-back*. The N-scan-back algorithm considers an ancestor hypothesis at time $k - N$ and looks ahead in time to all its children at the current time k (the leaf nodes). It evaluates the probabilities of all leaf nodes to find and keep the best branch at time $k - N$ and to discard all others.

V. Experiments and Results

The approach described above has been implemented and evaluated on a an ActiveMedia Powerbot mobile robot equipped with a Sick LMS laser scanner mounted at a height of 11 cm above ground. The angular resolution of the range scans was $0.5°$. Throughout all experiments we used the values listed in Table II for the conditional probabilities introduced in the previous section. Our adaptive method uses the probabilities with the superscript *F* for *free* tracks and the probabilities with the superscript *A* for *approved* tracks. We compare our method also to the non-adaptive case for which we use the probabilities with the superscript *F* as default values unless otherwise noted.

A. Person walking on an 8-shaped trajectory

In the first experiment a person follows a 8-shaped trajectory in a corridor of about 2.5 meters width in normal walking speed. As can be seen from Figure 1, our system was able to reliably track the person despite the fact that it

$p_{det\|F}$	$p_{occ\|F}$	$p_{del\|F}$	$p_{det\|A}$	$p_{occ\|A}$	$p_{del\|A}$	λ_{new}	λ_{fal}
0.3	0.63	0.07	0.2	0.79	0.01	0.001	0.003

TABLE II

PARAMETERS USED THROUGHOUT THE EXPERIMENTS.

only used a constant velocity motion model to track the sharp turns carried out by the person. The same leg tracks last over the entire duration of the experiment. This is illustrated by the diagram in right image of Figure 1 that shows a constant number of four tracks. Two of the four tracks are due to false alarms extracted in the clutter. Without adaptation of the occlusion probability, there is track loss at nearly every U-turn giving rise to many newly created tracks.

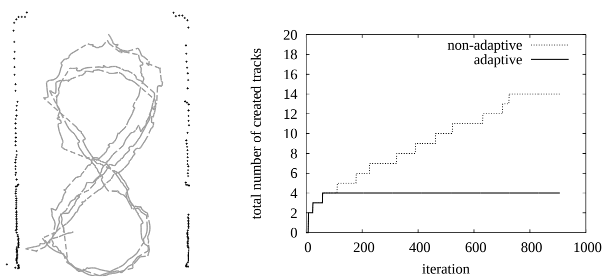

Fig. 1. Trajectories and total numbers of created tracks for experiment 1.

B. Person turning constantly while moving forward

In the second experiment a person is moving on a straight line turning $180°$ around the stationary leg at each step (see Figure 2). This unusual walking pattern produces heavy occlusions of the moving leg by the stationary one. The adaptive approach was able to track the person accurately during the experiment. The total number of tracks in Figure 2 (right) is constant (three), one of them being a false alarm. The mutual leg occlusion is poorly handled by the non-adaptive approach as the increasing number of new tracks in the diagram illustrates.

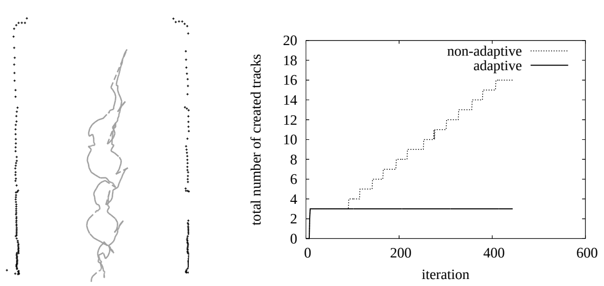

Fig. 2. Trajectories and total numbers of created tracks for experiment 2.

C. People walking randomly in a narrow corridor

It remains to be demonstrated that the superior performance of the adaptive approach found so far is not just due to better tuned probability parameters for approved tracks. This is demonstrated in the third experiment where up to four people simultaneously move through the field of view of the sensor. The subjects perform typical motion patterns at normal walking speed, they avoid each other, turn on the

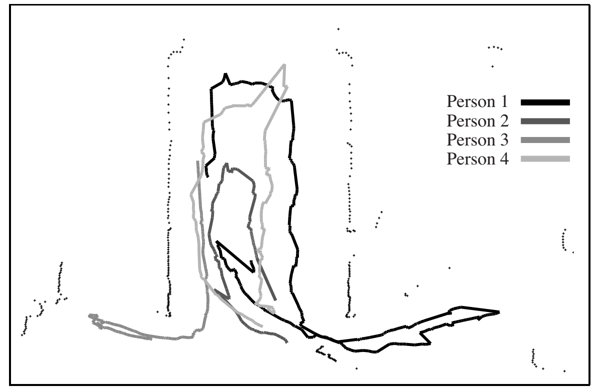

Fig. 3. Trajectories of four people tracks during experiment 3.

spot, cross paths, stop once in a while, and frequently enter and leave the field of view. This leads up to four validated people tracks simultaneously (eight leg tracks), not included false alarms due to, e.g., corners falsely detected as legs.

Figure 3 shows a portion of the experiment with four simultaneously tracked people. The chance of additional mutual occlusions from people is substantial in this narrow environment. Figure 4 depicts the total number of created tracks. Due to long lasting occlusions produced by other people, the system sometimes deletes tracks although the person is still there, and creates new tracks when the person becomes visible again. However, Figure 4 shows that compared to the non-adaptive case, we are able to track people more robustly over an extended period of time as the number of tracks is substantially closer to ground truth. The ground truth information was obtained by manual inspection.

If we use the parameter setting for approved tracks as default (and without adaptation), we observe in Figure 5 (left) that the number of simultaneous tracks nearly never decreases, that is, tracks are deleted with a very low probability. When tracks are not deleted, their uncertainty grows boundless producing a high level of ambiguity, and ergo, a high number of matching candidates that pass the Mahalanobis test. This causes an explosion of branches in the hypothesis tree as illustrated in Figure 5 (right). The diagram shows the number of hypotheses between steps 900 and 1000, the time when all four people were in the field of view. In the adaptive case, the peak numbers of hypotheses are seriously more moderate compared to the non-adaptive approach where the parameters for approved tracks are taken as default.

The average cycle time in this experiment with four people

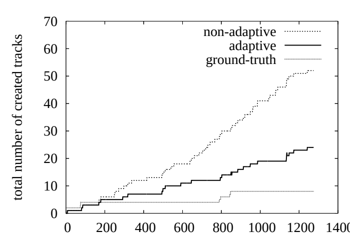

Fig. 4. Total number of created tracks for the adaptive method, the non-adaptive method, and the ground-truth.

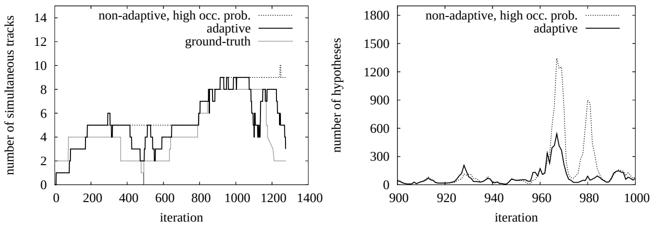

Fig. 5. Number of tracks for the adaptive method, the non-adaptive method with parameters for approved tracks as default versus the ground truth (left) and number of simultaneous hypotheses for our adaptive case and the non-adaptive method with parameters for approved tracks as default (right).

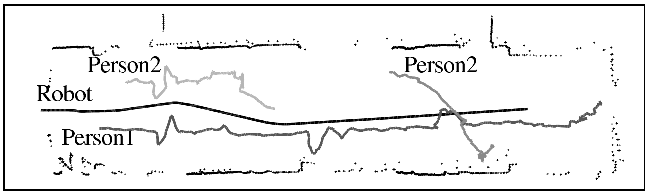

Fig. 6. Trajectories of robot and people in experiment 4. Person 1 is constantly tracked, person 2 receives a new identifier when reentering the sensor's field of view.

was 44.5 ms on an Athlon 4400+ and with a scan-back depth of eight (see section IV-B). A significant acceleration (from initially 220 ms) was due to the introduction of separate trees for tracks and hypotheses as proposed by Cox and Hingorani [13] that avoids processing duplicate tracks.

D. Tracking from a moving robot

In the forth experiment the robot moves with an average translational velocity of 0.33 m/s (max. 0.5 m/s) while tracking two people (Figure 6). The two subjects move at normal walking speeds, stop once in a while with person 2 leaving and re-entering the robot's field of view. Consecutive scans are aligned using odometry information. With a moving sensor, detection of moving leg blobs is more difficult as also the background becomes dynamic. Especially in clutter the AdaBoost classifier therefore generates a higher number of false alarms. Because people tracks are initialized only from leg tracks that satisfy our person model, the robot is able to robustly track the two people with only one incorrect people track that appears for two iterations. The non-adaptive approach creates additionally eleven incorrect leg tracks resulting in a total of four incorrect people tracks.

VI. Conclusions

In this paper, we addressed the problem of people tracking as a leg tracking problem utilizing an MHT for data association. We extended the original MHT to incorporate adaptive occlusion probabilities and present a mathematical derivation for this approach. The approach has been implemented and tested on a real robot with data acquired by a SICK laser range sensor. The experimental results demonstrate that our approach is able to robustly track multiple people based on observations of their legs even when enduring occlusions occur. We also carried out experiments that demonstrate that our adaptive approach outperforms a non-adaptive MHT with fixed occlusion probability settings, since it overly delays track deletion and thus produces a high level of ambiguity coupled with an explosion of the number of hypotheses. Our current system is able to perform each update fast enough for online processing on a state-of-the-art desktop computer even when the robot is tracking four people.

The occlusion model and the approach to extract people tracks have performed well in all our experiments. Still, they can both be replaced by more sophisticated models, independent of the theoretical results presented in this paper. Future work will aim at occlusion models for groups of people and a more rigorous technique to create people tracks from leg tracks.

VII. Acknowledgments

This work was partly funded by the German Federal Ministry of Education and Research (BMBF) under grant 01IME01F and by the European Comission under contract numbers FP6-IST-034120 and FP6-IST-045388.

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. Mataric, "Laser-based people tracking," in *Proc. of the IEEE Int. Conference on Robotics and Automation*, 2002.

[3] M. Kleinhagenbrock, S. Lang, J. Fritsch, F. Lömker, G. Fink, and G. Sagerer, "Person tracking with a mobile robot based on multi-modal anchoring," in *IEEE International Workshop on Robot and Human Interactive Communication (ROMAN)*, Berlin, Germany, 2002.

[4] D. Schulz, W. Burgard, D. Fox, and A. Cremers, "People tracking with a mobile robot using sample-based joint probabilistic data association filters," *International Journal of Robotics Research (IJRR)*, vol. 22, no. 2, pp. 99–116, 2003.

[5] E. Topp and H. Christensen, "Tracking for following and passing persons," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Alberta, Canada, 2005.

[6] 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.

[7] 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.

[8] D. Schulz, "A probabilistic exemplar approach to combine laser and vision for person tracking," in *Proc. Robotics: Science and Systems*, Philadelphia, USA, August 2006.

[9] M. Mucientes and W. Burgard, "Multiple hypothesis tracking of clusters of people," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Beijing, China, 2006.

[10] 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.

[11] J. Cui, H. Zha, H. Zhao, and R. Shibasaki, "Laser-based interacting people tracking using multi-level observations," in *IEEE/RSJ Int. Conference on Intelligent Robots and Systems*, Beijing, China, 2006.

[12] D. Reid, "An algorithm for tracking multiple targets," *IEEE Transactions on Automatic Control*, vol. 24, pp. 843–854, Dec 1979.

[13] I. J. Cox and S. L. 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.

[14] Y. Bar-Shalom and X.-R. Li, *Multitarget-Multisensor Tracking: Principles and Techniques*. Storrs, USA: YBS Publishing, 1995.

[15] K. O. Arras, Óscar Martínez Mozos, and W. Burgard, "Using boosted features for the detection of people in 2d range data," in *Proc. of the IEEE Int. Conference on Robotics and Automation*, Rome, Italy, 2007.

[16] K. G. Murty, "An algorithm for ranking all the assignments in order of increasing cost," *Operations Research*, vol. 16, pp. 682–687, 1968.