Adapting the State Uncertainties of Tracks to Environmental Constraints

Stephan Reuter and Klaus Dietmayer
Institute of Measurement, Control, and Microtechnology
University of Ulm, Germany
stephan.reuter@uni-ulm.de

Abstract – *In most multi-target tracking algorithms it is assumed that the movements of the targets are statistically independent of each other. This assumption may lead to predictions which are not possible due to physical exclusions. Instead of integrating the dependence between the objects directly into the tracking module, we propose to handle scenarios with interactions between the tracked object and other objects by adapting the uncertainty about the state of the object. The adaption is based on occupancy grids and reduces the uncertainty without endangering the consistency of the tracking filter.*

Keywords: Tracking, filtering, estimation, data association, extended objects.

1 Introduction

Most multi-target tracking algorithms like e.g. Multi Hypothesis Tracking (MHT)[1] or Joint Integrated Probabilistic Data Association (JIPDA)[2] use several single-target filters in order to track more than one target. Due to the use of single-target filters, a data association algorithm is necessary to assign the received measurements to existing tracks. In MHT a track hypothesis is created for every possible assignment between measurements and tracks. In order to keep the approach computationally tractable, a pruning procedure is necessary. JIPDA uses a probabilistic data association method which depends amongst others on the spatial measurement likelihood of the sensor. A possibility to avoid the data association problem could be the use of the Random Finite Set filter proposed by R. Mahler [6].

In scenarios with high object density this data association is one of the most challenging tasks due to ambiguity. The complexity of the data association depends on the process noise that is used in the prediction step of the filter, but in order to avoid inconsistency of the filter it is not recommended to choose a lower process noise in general.

An aspect which is often neglected in multi-target tracking is the influence of other objects in the environment on an object of interest. Thus, it is assumed that all targets move statistically independent of each other. Fig. 1 shows that it is possible to get physically impossible predictions of an objects state during occlusions, if we neglect the information about the environment.

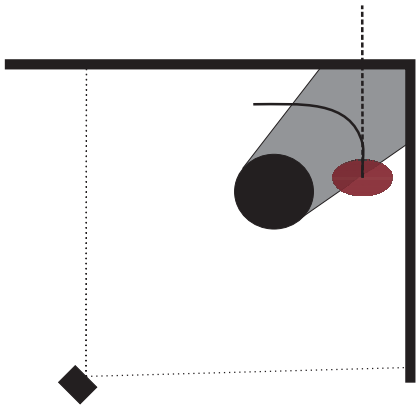

Figure 1: Standard prediction vs. expected prediction: constant velocity prediction is shown by a dashed line, solid line shows the prediction we would expect taking the obstacles into account. The gray colored area is occluded due to a static obstacle for the sensor located at the bottom left.

In this contribution an approach to integrate the social force model [5] into the prediction of occluded objects is introduced. This social force grid model represents the influence of static objects in the environment and is used to avoid physically impossible predictions. Further, the social force grid approach is extended in order to adapt the uncertainties about the state of the objects to the environmental constraints. The adapted uncertainties are only used for the data association, we do not change the uncertainties of the filter itself.

Amongst others, the social force grid is used to adapt the uncertainty to the environment by applying physical exclusions.

This contribution is organized as follows: First, the scenario and the sensor setup are introduced in section 2. Then, the social force grid and a prediction method based on this grid are introduced. In section 4 the new approach to adapt the uncertainty about the state of an object to the current environment is presented. The prediction of the adapted uncertainty to the next time step is introduced in section 5. Finally, tracking results are shown in section 6 to illustrate the gain which can be achieved by using the proposed approaches.

2 Scenario and Sensor Setup

In our demonstration scenario, a large room with a high pedestrian density is expected. Two laser range scanners are used to observe the room. Both scanners have a horizontal opening angle of 110° and an angular resolution of 0.25° and they are located in two corners of the room.

Due to the high density it is very likely that pedestrians are closely spaced, partly occluded or fully occluded for several measurement cycles. During occlusions the uncertainty about the state of a pedestrian is increasing with the number of measurement cycles where we do not have a measurement to update the pedestrians state. In Figure 2 the black circle illustrates the 3 σ bound of the uncertainty about the state of the occluded pedestrian a if we neglect environmental constraints. Depending on the data association method, it can be possible that the measurement of the yellow colored pedestrian b is associated to the track of the occluded person. In order to avoid an incorrect data association independent of the data association algorithm, it is necessary to integrate the environment into the state uncertainty. The blue dashed line shows the expected uncertainty of the state by taking the environment into account.

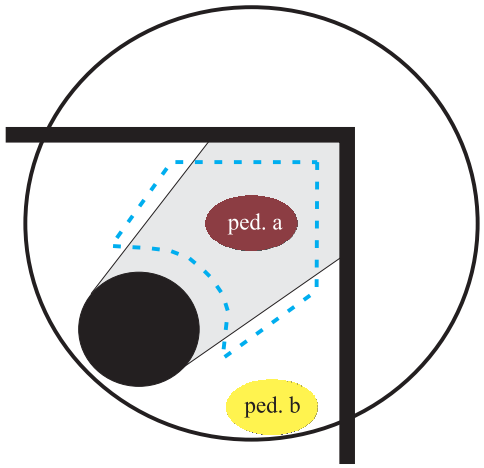

Figure 2: Standard uncertainty vs. expected uncertainty: uncertainty without environment adaption is shown by the black circle, the blue dashed line shows the expected uncertainty of the state of pedestrian a.

3 Social Force Grid

3.1 Social Force Model

In simulations for evacuation scenarios models representing the influence of the environment on individual pedestrians are used. One of these models is the *Social Force* model introduced by Helbing et. al [5] which is inspired by fluid dynamics. The social force model assumes that all objects around a person cause either attracting or repulsive forces. Further, it is assumed that there exists a destination point which the pedestrian wants to reach on the shortest path without collisions. A drawback of this model is that the force vectors may cancel out each other. Further, weighting the influence of the forces on the current moving direction is very critical.

Pellegrini et al. [3] proposed an extension of the social force model. They do not apply the forces directly to the motion vector of the pedestrian, but use the predicted point of closest approach for the driving force. This procedure corresponds to the anticipatory movement of pedestrians because the reaction to avoid a possible collision can be initiated earlier with this approach. The LTA model is mainly focused on the interaction between pedestrians. Pelligrini et al. compared the performance of their LTA model amongst others with the constant velocity model. Especially for scenarios with interacting people and short-time occlusions, the LTA model performs considerably better.

3.2 Social Force Grid

In order to be able to represent the influence of static objects independent of their shape, an occupancy grid is used in the proposed social force grid to represent the environment. Thus, even very complex shapes are easy to handle since we do not need to extract the shape. The repulsive forces of static objects can be modeled by an exponential function. Hence, we convolve the occupancy grid with a Gaussian kernel in order to include the forces. The grid cells of the occupancy grid, which are not observable, have to be declared as free space before the convolution is performed. Finally, we have to include the extension of the pedestrians into the grid. In order to reduce the computational costs it is expected that all pedestrians have the same size. The size of the pedestrians can be integrated into the grid by another convolution with a kernel that represents the size of a pedestrian.

In Fig. 3 the social force grid is shown. The center of the pedestrian may not be located at grid cells with a value of one due to physical exclusions. This social

force grid correponds to the environment shown in Fig. 1.

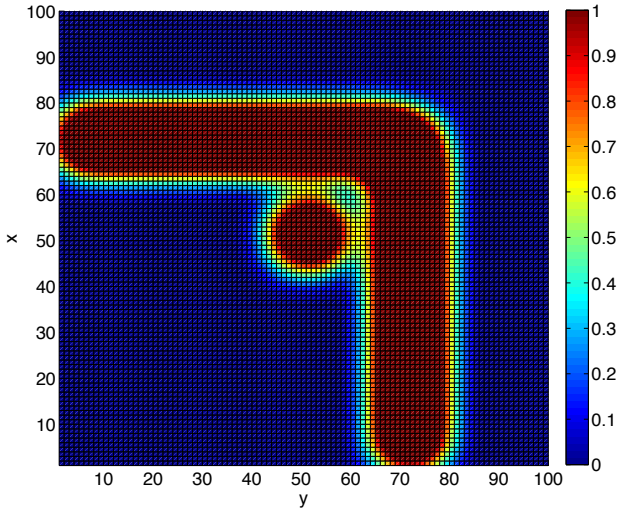

Figure 3: Social force grid calculated by the convolution of an occupancy grid representing the static objects shown in Figure 1 with a kernel representing the forces and the size of the pedestrian.

3.3 Prediction of Objects Using the Social Force Grid

The derived social force grid can be used to improve the prediction of occluded objects. The grid based prediction is done in the following way: if a pedestrian enters the occlusion area e.g. at position $(x, y) = (50, 62)$ with a velocity vector pointing in positive x direction, we predict the position of the pedestrian for n timesteps into the future according to Fig. 4. Since pedestrians

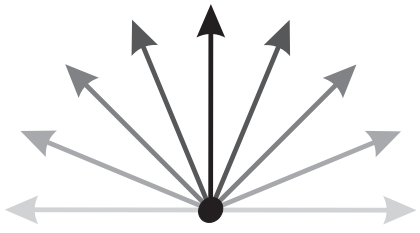

Figure 4: Path finding

normally try to reach their destination on the shortest path, they prefer to go straight on. Thus, the directions are weighted with a factor which represents this behavior. The behavior is modeled by the weights

$$w(\Delta\phi) = \left(\frac{1 + cos(\Delta\phi)}{2}\right)^{\alpha} \quad (1)$$

where α is a parameter for the peakiness of the function and $\Delta\phi$ is the difference between the estimated motion vector and the prediction. In our experiments, the best results are achieved with the parameter $\alpha = 4$. This function is similar to the one used in [3] to represent the influence of other objects on the movement of a pedestrian depending on the angle between the moving direction and the relative object position.

For each of the directions $\Delta\phi$ a cost function

$$J(\Delta\phi) = \frac{1}{w(\Delta\phi)} \cdot \sum_{i=1}^{n} g(x_{k+i|k+i-1}(\Delta\phi)) \quad (2)$$

is evaluated where $g(x_{k+i|k+i-1}(\Delta\phi))$ is the value of the social force grid at the grid cell corresponding to the predicted states $x_{k+i|k+i-1}(\Delta\phi)$ for the angle $\Delta\phi$. Further, the cost function is weighted with $\frac{1}{w(\Delta\phi)}$. Because

$$w(\Delta\phi) < 1 \;\forall\; \Delta\phi \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \setminus 0, \quad (3)$$

a change of the current moving direction leads to a multiplication of the costs with a factor greater than one. Finally, the most likely moving direction $\Delta\phi_{max}$ is found by searching the minimum of $J(\Delta\phi)$ for all $\Delta\phi$.

Because no measurement which belongs to the predicted object has been received and the behavior of the pedestrian and the prediction using the social force grid can be contradictory, the innovation step of the tracking filter is not calculated. Thus, the uncertainty about the objects position is increasing in the same way as if we would use the standard constant velocity (CV) model but the predicted position is different.

4 Adapting Uncertainty to Environmental Constraints

Using a Kalman filter, the uncertainty about the predicted position of an object is represented by the covariance matrix of a Gaussian distribution. Thus, the uncertainty can be adapted by weighting the probability density function (pdf) of the multivariate Gaussian distribution.

In order to adapt the uncertainty of the predicted state of an pedestrian, several aspects have to be taken into account: the possible locations of the pedestrian due to static objects, the spatial detection probability of the sensor(s), the possible re-appearance positions, and the positions of other pedestrians. Depending on the current situation, we do not have to consider all of these aspects, e.g. the spatial detection probability and the re-appearance positions are only necessary for occluded pedestrians.

The influence of static objects can be modeled by the social force grid presented in the previous section. Thus, for every grid cell c_i the probability that the center of the object may be located in the cell is given by

$$p_s(c_i) = 1 - g(c_i) \quad (4)$$

where $g(c_i)$ is the value of the social force grid of cell c_i.

The spatial detection probability can be calculated by the measurement grid which is used to calculate the occupancy grid. Each cell of the measurement grid is initialized with the value 0.5. The value of all cells between the location of the sensor and a measurement, which are assumed to be free space, are decreased while the value of cells which contain measurements are increased. Thus, all grid cells which can not be observed by the sensor remain at the initialization value. Consequently, the binary spatial detection probability is given by

$$p_d(c_i) = \begin{cases} 1 & x \neq 0.5 \\ 0 & x = 0.5 \end{cases} \quad (5)$$

In Fig. 5 the grid map representing the detection probability is shown.

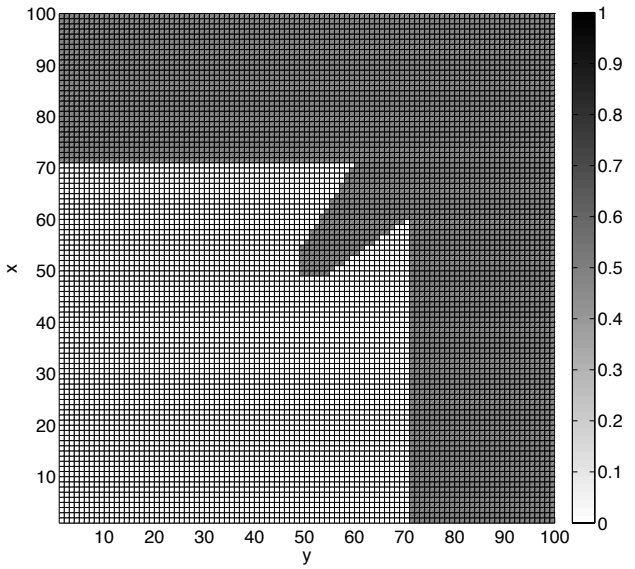

Figure 5: Grid map representing the detection probabilities. White colored grid cells are observable for the sensor located at $(0, 0)$, gray colored cells are occluded due to objects between sensor and the grid cell.

By using the spatial detection probability and the social force grid, we can limit the uncertainty to the occluded area and the possible locations. In order to be able to represent the case that a pedestrian re-appears, this area has to be extended. Therefore, a grid which represents the possible re-appearance locations is used. Since the scan points are ordered by angle, possible re-appearance locations are given by the line between two succeeding points whose radial distance exceeds a threshold. Depending on the measurement rate, we have to extend the re-appearance area a certain distance into the observable area of our sensor. Further, the distance has to depend on the detection probability of the sensor. This extension can be approximated by a convolution of the grid containing the lines with a kernel function, e.g. a Gaussian kernel. In Fig. 6 the re-appearance map for our example is shown.

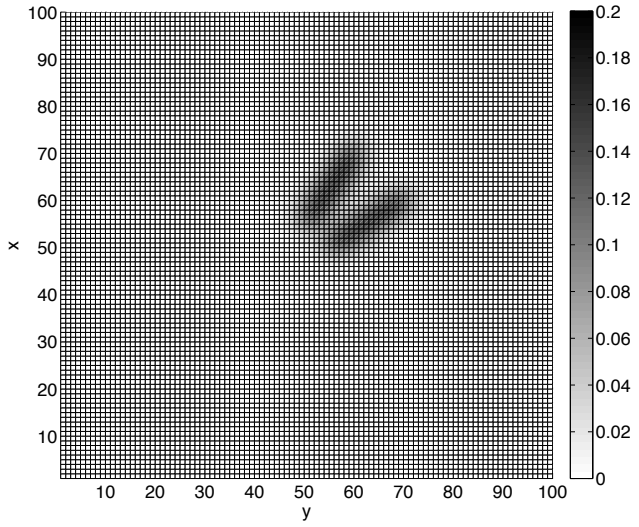

Figure 6: Grid map which represents the re-appearance areas for the occluded pedestrian.

Next, we need a further grid map which represents other pedestrians in the environment. In order to represent them, we use an approach similar to the one presented in [7]. The extension of an object can be incorporated by convolving the uncertainty about the position of a point target with a two dimensional elliptical kernel function $E_\phi(L_i, W_i)$ which may be rotated by an angle ϕ. The lenght L_i and width W_i correspond to the principle axes of the ellipse. Thus, the extended object occupancy likelihood (EOOL) of each pedestrian is represented by

$$D_i(s) = \left(\mathcal{N}(s, \hat{\mathbf{x}}_i, \mathbf{P}_i)_{k|k} \cdot p(\exists x_i)\right) * E_\phi(L_i, W_i), \quad (6)$$

where $*$ denotes the two dimensional convolution. $\mathcal{N}(s, \hat{\mathbf{x}}_i, \mathbf{P}_i)_{k|k}$ is the Gaussian distribution of the uncertainty about the position of the point target and $p(\exists x_i)$ is the existence probability of the track. The EOOL is also approximated by a grid. Since the EOOLs of several pedestrians are not independent of each other, we approximate the resulting EOOL by

$$p_{ped}(c_i) = \max_j(D_j(c_i)) \quad (7)$$

In Fig. 7 a grid map representing three pedestrians is shown. The pedestrians have different sizes, different state uncertainties, and existence probabilities.

Finally, we have to combine the information of all grids to be able to adapt the uncertainty about the position of the object. The adapted uncertainty for an

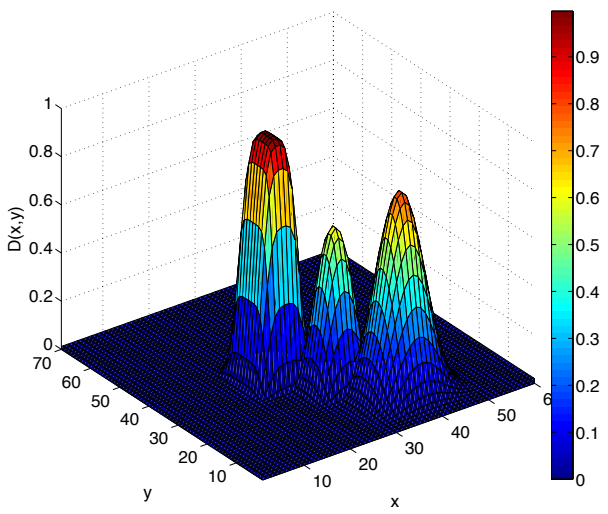

Figure 7: Extended object occupancy likelihood of three pedestrians with different state uncertainty and existence probability.

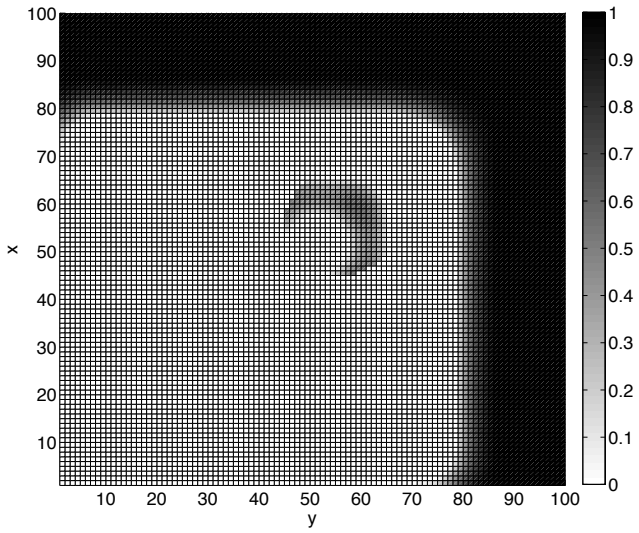

Figure 8: Adaption grid map to weight the state uncertainty of the occluded pedestrian. Other pedestrians in the environment are neglected here.

occluded pedestrian is calculated by

$$u(c_i) = f(c_i) \cdot p_s(c_i) \cdot sgn(p_d(c_i) + p_r(c_i)) \cdot (1 - p_{ped}(c_i)) \tag{8}$$

where $f(c_i)$ is the discretized Gaussian distribution representing the uncertainty about the state of the object and $p_r(c_i)$ is the value of the re-appearance grid at grid cell c_i.

In Fig. 8 the grid map to adapt the uncertainty about the state of a pedestrian is shown prior to the multiplication with the Gaussian distribution $f(c_i)$. Further, the influence of other pedestrians in the environment is neglected in this figure. We observe that the adaption grid map looks similar to the uncertainty which is expected if we take the environment into account (see Fig. 2).

5 Prediction of the Adapted Uncertainty

In the constant velocity model with state vector $\underline{x} = (x, \dot{x}, y, \dot{y})$ the a priori state is calculated by

$$\mathbf{x}_{k+1|k} = \underbrace{\begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\mathbf{F}} \cdot \mathbf{x}_{k|k}, \tag{9}$$

and the a priori covariance of the estimation is given by

$$\mathbf{P}_{k+1|k} = \mathbf{F} \cdot \mathbf{P}_{k|k} \cdot \mathbf{F}^T + \mathbf{Q}, \tag{10}$$

where $\mathbf{Q}$ is the process noise of the motion model.

If the a priori covariance of the estimation is very high due to several missed detections, the presented adaption method may not deliver correct results by adapting the a priori covariance directly. Thus, a prediction of the adapted uncertainty to the next time step has to be performed.

Due to the adaption, the uncertainty is not Gaussian any more. Thus, another possiblity to predict the uncertainty is necessary. The first possibility is to approximate the adapted uncertainty with a Gaussian distribution which can e.g. be calculated by using the EM-Algorithm [8].

Depending on the shape of the adapted uncertainty, this approximation might be quite inaccurate. Thus, we propose to perform the prediction directly with the grid. Anyway, the prediction of the adapted uncertainty has to be done according to (10). The additional uncertainty we get by predicting the state is given by

$$\mathbf{Q}_F = \mathbf{P}_{k+1|k} - \mathbf{P}_{k|k}. \tag{11}$$

Since the convolution of two Gaussian distributions

$$g_1(x) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{\sigma_1^2}} \tag{12}$$

and

$$g_2(x) = \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(x-\mu_2)^2}{\sigma_2^2}} \tag{13}$$

with each other leads to

$$h(x) = g_1(x) * g_2(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{\sigma^2}} \tag{14}$$

with

$$\mu = \mu_1 + \mu_2 \quad (15)$$
$$\sigma^2 = \sigma_1^2 + \sigma_2^2, \quad (16)$$

we can represent the prediction by a convolution with a Gaussian kernel with mean value $\mu = \mathbf{x}_{k+1|k} - \mathbf{x}_{k|k}$ and covariance $\Sigma = \mathbf{Q}_F$. Since we use only a two dimensional grid map, only the components representing the position are used for the convolution.

6 Results

The presented methods to improve the prediction and adapt the uncertainty are compared to a standard Kalman filter in this section. We use real sensor data of a scenario similar to the one shown in Fig. 2 in this section. Due to the occlusion, some parts of the wall at $x \approx 3.6$ are not visible in the figures. A standard constant velocity Kalman filter is used to track the pedestrian while it is not occluded. If the pedestrian enters the occlusion, the prediction of the state of the pedestrian is calculated using the social force grid. Further, the grid map to adapt the uncertainty is only calculated during occlusions. The grid map is only used to simplify the data association, it does not change the state covariance of the tracker.

In Fig. 9 the predicted states of a pedestrian are shown for a standard CV tracking filter and a filter using the social force grid method. We observe that the CV filter predicts the pedestrian straight on while the social force grid method initiates the pedestrian to turn left. Although the pedestrian is only occluded for less than one second, we observe that the re-appearing pedestrian is outside the 3 σ gate of the constant velocity tracker. On the other hand, the predicted position of the pedestrian very close to the predicted position based on the social force grid.

In the following the normalized innovation squared (NIS) error [9] is used to evaluate the performance of the prediction using a social force grid. The NIS is calculated by

$$\epsilon(k) = \left[z_k - \hat{z}_{k|k-1}\right]^T \cdot S_k^{-1} \left[z_k - \hat{z}_{k|k-1}\right] \quad (17)$$

Since the NIS value is defined as the residual between measurement and estimated position we can not calculate it if we do not receive a measurement. Therefore, we set the NIS value to minus one in this case. In Fig. 10 we observe that the NIS value of the tracker using the social force grid is much smaller than the value of the one using the constant velocity model at the measurement steps no. 304 and 360 where the pedestrian re-appears after occlusions for about 10 measurement cycles (i.e. 800ms). Further, due to the smaller NIS error at the re-appearance time of the pedestrian, the error of the social force grid method is also smaller for the succeeding measurement cycles. It takes about 15

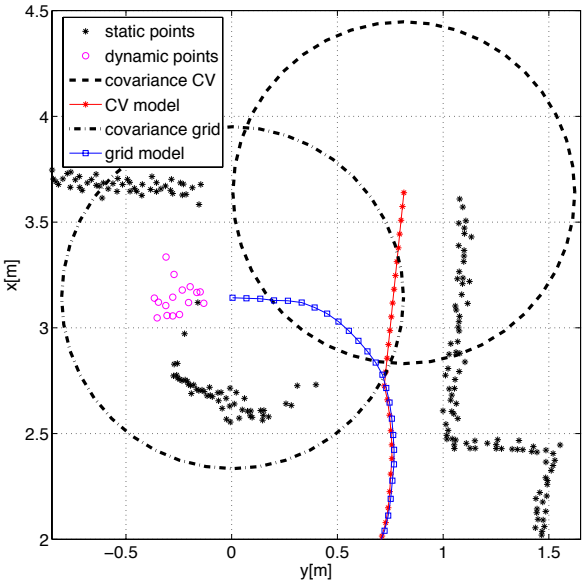

Figure 9: Estimated Positions of the pedestrian for the CV model (red) and the social force grid model (blue). Due to the occlusion, we do not observe parts of the wall at $x \approx 3.6$m. Further, the 3 σ ellipses are shown for the time when the pedestrian re-appears.

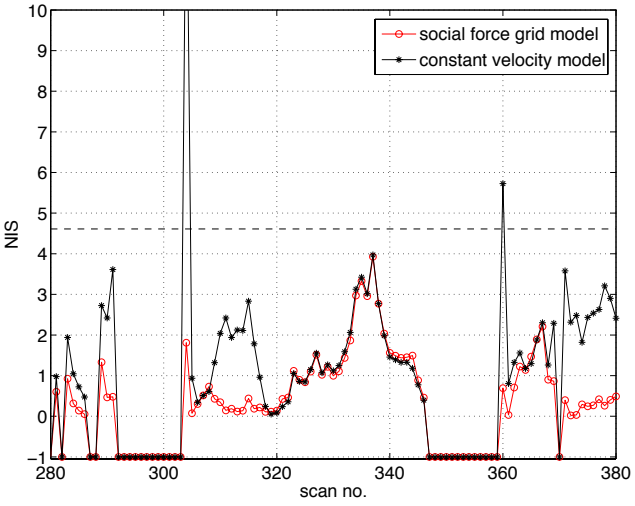

Figure 10: Normalized innovation squared for the CV motion model and the social force motion model.

measurement cycles (1.2 seconds) until the NIS error of the constant velocity model converges again to the value of the social force grid model. At measurement no. 304 we would further not associate the measurement with the existing track of the CV tracker if a 3 σ bound is used for gating. Thus, a new track would be initialized due to inconsistency of the CV tracker. Consequently, depending on the initialization procedure it would take some time until the track is confirmed again.

In Fig. 11 the adapted uncertainty about the state of the object is illustrated by red circles. We observe that the adapted uncertainty is much smaller than the uncertainty of the filter itself. Due to the environmental constraints, the adapted uncertainty does not enlarge any more, even if the pedestrian stays in the occluded area for a very long time. On the other hand, the state uncertainty of the filter increases without a limit. Consequently, measurements of other tracks could be associated to this track.

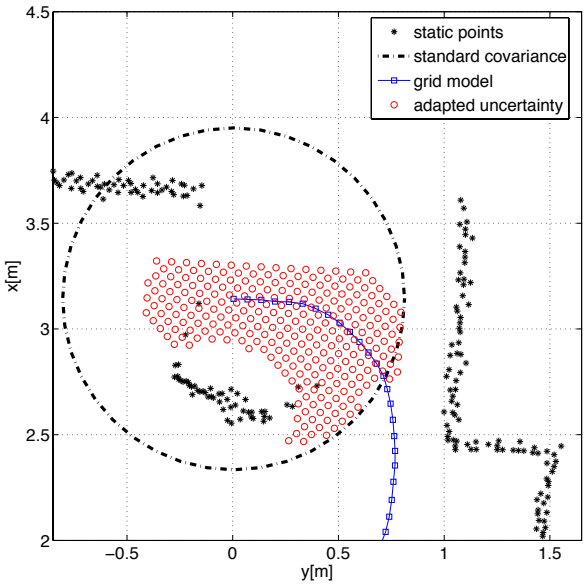

Figure 11: Adapted uncertainty vs. standard covariance: red circles mark the area of the adapted uncertainty, the black circle shows the 3 σ bound of the covariance matrix.

Considering only the adaption of the uncertainties due to static obstacles, we can save computational costs by calculating the adaption grid shown in Fig. 8 only if the environment changes. Using the EOOL approach introduced in section 4 we obtain approximately the same results as shown in Fig. 11, if the tracked pedestrian is occluded by another pedestrian instead of a static obstacle.

7 Conclusion

In this contribution a method to improve the prediction of occluded objects based on a social force grid has been introduced. Further, the social force grid was extended to integrate environmental constraints in the data association step by adapting the uncertainty about the state of an object. The uncertainties are reduced by the proposed method without endangering the consistency of the tracking algorithm. This leads to a simplified data association because a smaller number of measurements can be associated to a track. Especially in scenarios with long-time occlusions the proposed method outperforms standard approaches since the uncertainty about the state of an object is limited due to environmental constraints.

In future we plan to integrate an existence estimation for the occluded objects, since otherwise objects will never be deleted if they do not leave an occlusion again.

8 ACKNOWLEDGMENTS

This work is done within the Transregional Collaborative Research Center SFB/TRR 62 "Companion-Technology for Cognitive Technical Systems" funded by the German Research Foundation (DFG).

References

[1] S. Blackman and R. Popoli, "Design and Analysis of Modern Tracking Systems", *Artech House*, Boston, 1999

[2] D. Musicki and R. Evans, "Joint Integrated Probabilistic Data Association: JIPDA", *IEEE Transactions on Aerospace and Electronic Systems*, vol. 40, no. 3, July 2004

[3] S. Pellegrini and A. Ess and K. Schindler and L. van Gool, "You'll never walk alone: modeling social behavior for multi-target tracking",*IEEE Int'l Conf. on Computer Vision (ICCV)*, Kyoto, Japan, 2009.

[4] T. Weiss, B. Schiele, and K. Dietmayer, "Robust Driving path Detection in Urban and Highway Scenarios Using a Laser Scanner and Online Occupancy Grids", *IEEE Intelligent Vehicles Symposium 2007*, Istanbul, Turkey, June 13-15, 2007, pp. 184–189.

[5] D. Helbing and P. Molnár. "Social force model for pedestrian dynamics", *Physical Review E*, 51(5):pp. 4282–4286, 1995.

[6] R. Mahler, "Statistical Multisource-Multitarget Information Fusion", *Artech House*, Boston, 2007

[7] M. Munz, M. Mählisch, J. Dickmann, K. Dietmayer, "Probabilistic Modeling of Sensor Properties in Generic Fusion Systems for Modern Driver Assistance Systems", *IEEE Intelligent Vehicles Symposium 2010*, San Diego, USA, June 21-24

[8] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm", *Journal of the Royal Statistical Society, series B*, vol. 39, pp. 1–38, 1977.

[9] Y. Bar-Shalom, X. Rong Li, and T. Kirubarajan, "Estimation with Applications to Tracking and Navigation", *Wiley*, New York, 2001.