In Proceedings of 9th Workshop on Humanoid Soccer Robots
14th IEEE-RAS International Conference on Humanoid Robots (Humanoids), Madrid, Spain, 2014.

Multi-Hypothesis Goal Modeling for a Humanoid Soccer Robot

Marcus M. Scheunemann and Heinrich Mellmann
Department of Computer Science Artificial Intelligence Laboratory
Humboldt-Universität zu Berlin
Unter den Linden 6, 10099 Berlin, Germany
{scheunem,mellmann}@informatik.hu-berlin.de
http://www.naoth.de

***Abstract*—Information about objects and their positions in an environment are necessary requirements for most tasks of a mobile autonomous robot, in particular regarding the control of behavior and navigation. This presents a specialchallenge for robots with a limited view angle.**

Autonomously soccer playing robots in the dynamic environment of the RoboCup Standard Platform League are exposed to these difficulties. Most approaches aggregate all available information in one holistic model in order to localize robots. In case of inconsistent perceptions the model either turns noisy or creates and tracks an addiotional hypothesis. To improve the localization – and thus the behavior control – local models have received only little attention so far.

In this work the implementation of a local goal model is presented and analyzed. A multi-hypothesis particle filter is used to cope with ambiguity of goal post percepts as well as to process incomplete and uncertain sensor information. Additionally, a percept buffer supports the initialization and also facilitates the handling of sparse false measurements. On the basis of this local goal model inconsistencies can be explicitly modeled, which may be used to stabilize the location of a robot.

I. INTRODUCTION

Many tasks of a mobile robot, e.g., navigation, require the knowledge of the positions of surrounding objects. This is especially challenging for robots whose perception is based on a directed visual system, e.g., a camera with limited view angle. The incomplete and noisy sensor information leads to uncertainty in the robots' belief of the world. An appropriate world model is necessary to enable the robot to make plans and to achieve complex behavior.

Most current approaches are based on strong assumptions regarding the capabilities of visual perception and the kinematic structure of the robots. Those assumptions are considered common knowledge and are barely discussed explicitly. Within RoboCup the spatial situation modeling is mainly based on visual perception. Visual processing is considered a noisy black box sensor perceiving objects like goal posts, ball, etc. To estimate the relation between visual perceptions and positions of the perceived objects the kinematics of the robot is assumed to be known.

In reality the visual processing has a very complex and often not deterministic detection behavior. This is due to the fact that the actual environment is dynamic and may have unknown aspects as for instance the light conditions may change or the field might be surrounded by a crowd with colorful clothes. Similarly, the assumptions about the kinematic structure of the robot are violated in reality. The individual robots exhibit often certain deviations from default specifications due to production inaccuracies or simply abrasion over time.

Those violations may result in inconsistent observations. For instance, false positive visual perceptions, e.g., not existing goal posts, or a slightly shifted camera resulting in a incorrect projection of a detected object. Dealing with inconsistent observations is a central element of a modeling algorithm. An inconsistency can occur because the model does not represent the actual state correctly or because the observation itself is not correct, e.g., a false positive. In the global probabilistic localization approaches inconsistencies are resolved implicitly by increasing the likelihood of the positions confirming this observation. For example, in case of a goal post the possible positions result in a circle around it. Thus, incorrect observations may lead to very complicated likelihood. Since in reality representational power is limited due to limited resources, this might cause major issues and even lead to degeneration of the model. The common approach is to reduce those violations in both cases by careful parameter calibration for a given scenario, e.g., adjusting color parameters at the competition site or calibrating the joint offsets for every particular robot.

In this work we take a closer look at the specific nature of spatial modeling based on visual perception. We investigate the structure of errors in sensory input and present a hierarchical modeling approach based on multi-hypothesis object tracking targeting specifically those errors. In our experiments we apply this approach to model the soccer goal in the SPL context.

The reminder of the paper is structured as follows. In the next section we discuss sources of different kinds of uncertainty and their influence on spatial modeling. In section III the Multi-Hypothesis Goal Model (MHGM) is introduced. Experiments are presented and discussed in section IV. Section V summarizes and concludes the overall work.

II. SPATIAL MODELING BASED ON VISUAL PERCEPTION

In general, sensory information is noisy and incomplete, leading to uncertainty in the robots' belief of the current situation. Usually, this uncertainty is modeled homogeneously, e.g.,

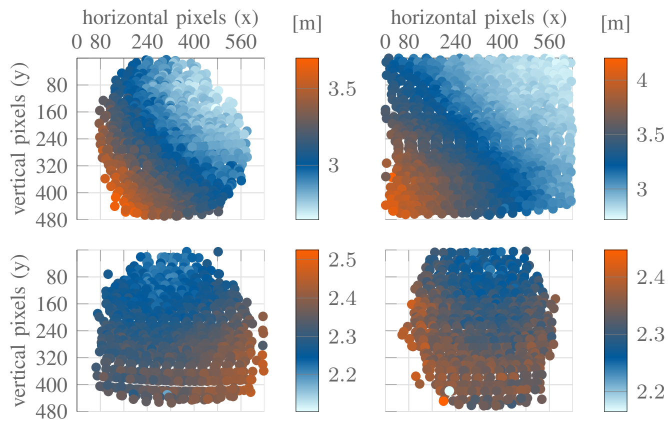

Fig. 1. Estimated distance to a goal post depending on its position in the image. The actual distance is 3m. The robot moves only its head to ensure the goal post is detected accurately, the head speed is very slow. Each point represents a goal post detection in corresponding image coordinates with the color illustrating the estimated distance to it. The results for four different robots are shown.

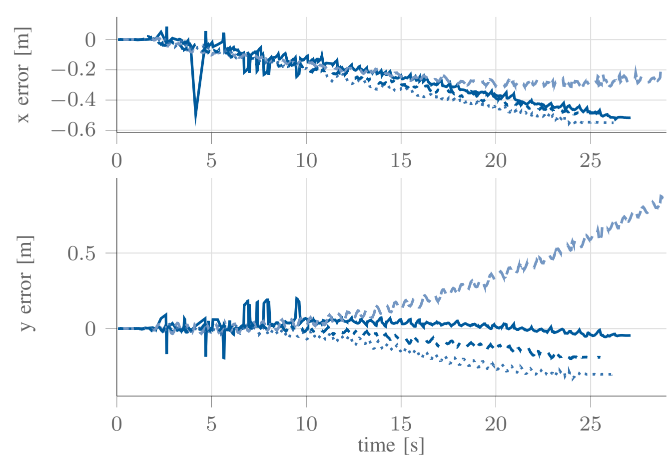

Fig. 2. In this experiment, four robots are forced to walk a straight line. The x and y components of the path traveled by each robot (odometry) are plotted separately over time.

as a probability distribution in probabilistic approaches. However, on closer inspection we can identify three qualitatively different components: *noise*, *false detections* (false positives) and *ambiguity*. The perception noise can be originated by discretization of the image processing and kinematic model. False positives are wrong perceptions which do not correspond to any known object, e.g., goal post detected in the colored clothes of the audience. Usually, such wrong perceptions occur sporadically (sparse false positives), but also systematic occurrences are possible (dense false positives). An example for dense false positives can be yellow pants of a visitor detected as a goal post. Finally, ambiguity results from the fact, that sensors are not able to observe the whole state of the robot, i.e., a particular observation has multiple interpretations. A goal post observed by the robot yields a circle of possible robot' positions around it. Therefore, the task of the modeling algorithm is to integrate the partial observations and to retrieve the state of the robot. In the presented approach we investigate how those aspects of uncertainty can be treated separately, which may lead to much more robust modeling approaches.

In a typical molding approach within RoboCup uncertainty is mainly originated from two different sources: image processing and kinematic model. The uncertainty of visual perception has already been discussed briefly in the introduction. A special problem here is caused by the false positive detections, e.g., pants detected as goal posts outside of the soccer field. A usual approach to tackle this issue is a tedious calibration of the vision algorithm for a given environment. For the kinematic model it is very common, to assume the specifications of the robot's body as given, e.g., masses of the links or lengths of the limbs, since those are usually known from the construction process. Based on these assumptions and proprioceptive sensors like accelerometer, gyro and FSR it is possible to derive a model of the robot's kinematics. This model can be used to estimate the transformation between image coordinates and local robot coordinates, which is also called *camera matrix*. Additionally, this kinematic model can be used to track the motion of the robot and provide an estimation for the relative traveled path, commonly referred to as odometry. The assumptions concerning the kinematics provide, in general, an acceptable approximation for the most use cases. The individual robots, however, exhibit often certain deviations from those specifications due to production inaccuracies or simply abrasion over time. Figure 1 and 2 visualize the deviations across different robots. A popular and most direct approach here is to calibrate kinematic parameters individually for each particular robot. Obviously, such calibrations have to be performed regularly based on the usage intensity of the robot. Most recent examples for such approaches are presented in [1].

Basically, the global localization problem is to find the transformation between the local coordinates of the robot and a certain global coordinate system, e.g., the field coordinates. Thereby, in a global approach the state space is modeled in the global coordinates, which means the transformation from the local into the global coordinates is solved. The combined uncertainty, caused by noise, false positives and ambiguity, from the different sensors accumulate and may take a very complicated shape in the global coordinates. In particular, it can be amplified as discussed in [2].

Based on the assumption of a known environment, the modeling of static elements like goals or penalty area is reduced to determining the position of the robot relative to a certain global coordinate system, i.e., localize the robot within a given map. Currently, the most successful approaches are based on the probabilistic Bayes-filtering. Here it is tried to solve the global localization problem by integrating the sensory data into a probability distribution describing the position of the robot on the field. In particular, the particle filter [3] and multi-hypothesis Kalman filter [4] are currently the most popular implementations.

Particle filters are easy to implement and are able to represent complex distributions. In general, it requires a large number of particles to represent the probability distribution adequately. In the most implementations, however, the number of particles is chosen very low due to limited computational resources. Essentially, this leads to a filter which is tracking a maximum of the underlying probability distribution.

Multi-hypothesis approaches represent the global probability distribution by a number of unimodal hypotheses, usually a mixture of gaussians. Here a new hypothesis is introduced each time an observation is inconstant with the current model. Those hypotheses are maintained in parallel until the inconsistency can be resolved and less likely hypotheses are deleted or merged. In [4] new hypotheses are introduced by splitting the old ones for each ambiguous observation and in [5] by creating new hypotheses based on strong sensory data similar to sensor resetting in particle filters.

While recent self localization solutions seem to work out well in the RoboCup environment, they still exhibit strong sensitivity regarding inconsistencies caused by the noise from the kinematic model and visual system, whereby false positives are specially problematic. The combined uncertainty caused by noise, false positives and ambiguity from the different sensors accumulate and may take a very complicated shape in the global coordinates. In particular, it can be amplified as discussed in [2]. Local modeling approaches and separate filtering of different components of uncertainty may lead to more robust solutions. Here, local multi-hypothesis object tracking approaches provide a strong basis. They are well studied in related fields and have proven to be very effective, e.g., [6], but so far gained comparably little attention within the RoboCup community, e.g., [7].

III. Multi-Hypothesis Goal Model (MHGM)

In this section we describe a multi-hypothesis approach for modeling a soccer goal within the RoboCup context. The goal is an essential feature in a soccer game and provides enough complexity to verify the ideas discussed in the previous section. The whole goal is rarely observed and we assume the image processing to detect separate goal posts. So we represent the goal by its corresponding posts. To reduce complexity of the shape of uncertainty we model the separate goal posts in local robot coordinates. The ambiguous goal posts are tracked by a multi-hypothesis particle filter. The actual goal model is extracted from the set of post hypotheses.

As discussed in the previous section, the joint uncertainty can be subdivided in noise, false detections and ambiguity. Each of this components is treated separately in our approach. The multi-hypothesis filter has to take care of noise and false detections, but it does not resolve the ambiguity of the goal posts. Instead, all occurring goal posts are represented by corresponding hypotheses and the ambiguity is solved on the next level when the goal model is extracted. Particle filters are great in filtering noise and are shown to be very effective for object tracking. To deal with sparse false positives we introduce a delayed initialization procedure. We assume a false positive to result in an inconsistency, i.e., it cannot be confirmed by any existing goal post hypothesis. In this case the percept is stored in a short time buffer for later consideration. This buffer is checked for clusters, in case a significant cluster of goal post percepts accumulated during a short period of time, a new hypothesis is initialized based on this cluster. The dense false detections result in post hypotheses, which is later ignored while extracting the goal.

In the following we describe the algorithm in more detail. Please note, the basics and the according theoretical background for particle filters as well as multi-hypothesis approaches can be found in [6], [8].

A goal post percept can be described as a vector $(\alpha, d) \in \mathcal{X} := [-\pi, \pi] \times \mathbb{R}_+$ in the local polar coordinates of the robot, where d is the distance to the goal post and α the horizontal bearing angle. The model (hypothesis)of a single goal post is represented by a final set of particles $H \subset \mathcal{X}$. The whole model consists of a set of post hypotheses $\mathcal{H} := \{H_1, ..., H_n | H_i \subset \mathcal{X}\}$ and a percept buffer $B \subset \mathcal{X}$.

Algorithm

The Algorithm is triggered by a new image. As input data we assume a set of detected post percepts $P = \{p_1, ..., p_k\} \subset \mathcal{X}$ and odometry information $o \in SE(2)$, i.e., a 2D coordinate transformation describing robot's motion since the last update. In each cycle we execute the following steps.

1) **Prediction:** Update all particle filters in $\mathcal{H}$ and the buffer B by odometry information, i.e., apply the transformation o to each element of the sets;
2) **Update**
    a) Association: solve the stable marriage problem between the post hypotheses $\mathcal{H}$ and the perceptions P. Note that, in general, those sets may have a different number of elements. Update all hypotheses $H \in \mathcal{H}$ with the corresponding match $p \in P$, if the weight of the match is higher than a given threshold. Add the remaining observations to the buffer B.
    b) Resample all particle filters $H \in \mathcal{H}$.
3) **Manage:**
    a) Remove all filters $H \in \mathcal{H}$ whose *weight* is below a defined threshold. Thereby, the weight of a filter is calculated by a weighted integral of the prior weightings;
    b) Remove old percepts from the buffer B;
    c) Cluster the buffer by euclidean distance. If there is a *sufficiently large* cluster $C \subset B$, use C to initialize a new filter $H \subset \mathcal{X}$ and add it to $\mathcal{H}$;
4) **Model extraction:** Find $H_1, H_2 \in \mathcal{H}$, so that $d := (\|E(H_1) - E(H_2)\|_2 - goalWidth)^2$ is minimized. If d is smaller than a specific threshold, sort both candidates under the assumption that the robot is located inside the soccer field. Without loss of generality, set $p_L = E(H_1), p_R = E(H_2)$ as positions of the left and right post of the goal model. Hereby, $E(H_i)$ is the expected value of the corresponding hypothesis, i.e., the mean of the particle filter.

Remarks

The minimal cluster size required for the **initialization** of a new hypothesis controls the delay in the initialization and the robustness regarding sparse false positives. **Association** problem between the percepts and hypotheses is resolved explicitly, thereby it is done with a one by one assumption. Usually, one post generates only one percept, so that the assumption is fulfilled. Hypotheses are **deleted** when their weight is decreased. This is done when the corresponding particle filter was not updated by a percept for a certain time. Therefore, there is no need to merge filters. When two filters are representing the same object, one of them will not receive any updates and will be eventually removed. The current approach to **extract the model** is quite simple, but has shown good results in our experiments.

IV. Experiments

In order to test the performance of our modeling approach we performed experiments on a real robot. Two of them are represented in this paper. In the first experiment we investigate the robustness of the goal model regarding dense and sparse false positive perception. The second experiment aims to test the qualitative ability to keep track of the goal model in a dynamic situation. Since MHGM is a randomized algorithm, the results may vary even if applied to the exact same input data. To exclude errors in our tests, MHGM is tested on recorded perception which allows for repeated executions on the same input data.

A. Sparse and Dense False Positives

The aim of this experiment is to investigate the robustness of the MHGM regarding false positives and uncertain perceptions in a controlled environment. Figure 3 (left) shows the experimental set up. The robot is aligned in the center of the field facing a true soccer goal. An additional post is added to the scenario. From the object recognition perspective, this goal post is identically to the other posts. This simulates dense false positives caused by post-like structures, e.g., colored pants, which may occur in the environment of a regular RoboCup game. During the experiment, the robot moves only its head in a simple circular search motion. In this setup, the goal detection is not very hard and no sparse false positives occur. To simulate this, random artificial goal post observations are generated and added to the process. A maximum of one false detection is generated with a probability ρ in each frame. The observations are uniformly distributed within a relevant region of the image where the goal is typically perceived. Those percepts are seamlessly added to the modeling process together with the regular percepts. Figure 3 (center) shows all the percepts collected during an experiment with $\rho = 0.1$. The color indicates the classification of the percepts by the MHGM. A systematic distortion of perception and the resulting model can be clearly observed. This distortion is caused by the uncalibrated kinematic model. In Figure 3 (right) one can see a snapshot of the model at the end of the experiment. All three goal posts were modeled and classified correctly by corresponding particle filters and the goal model was extracted as expected.

Fig. 4 visualizes the association of percepts by the MHGM with $\rho = 0.1$. It can be clearly seen, that except for a single outlier all percepts were classified correctly. In particular, all dense false detections caused by the additional goal post were modeled in a separate hypothesis (dark gray) and did not affect the resulting goal model. All sparse false positives are sorted in the buffer (indicated by light gray color) except for one at time 9.66. The percept was switched with the perception of the right goal post (blue). This is not necessarily a sign of wrong model calculation, since sparse false perceptions are distributed uniformly, it can happen that a single false positive is closer at the goal post than the actual detection.

During the search motion the robot moves its head from left to right and vice versa. Thereby it looks more up when moving the head from left to right to inspect the distant areas and more down when moving from right to left. The motion from right to left is faster than the one in the opposite direction. As can be easily seen in Figure 4, this results in long periods of stable post detection when moving from left to right and rather short and more sporadic detections when moving from right to left.

TABLE I
Incorrect associations of false positive goal post percepts.

Probability of a false positive ρ	0.1	0.3	0.5	0.7	1.0
Wrong associations mean	1.4	5.2	7.6	11.2	46.4
Wrong associations standard deviation	1.3	2.0	1.5	4.5	39.4

The whole experiment was systematically repeated with different probabilities ρ for false positives with five repetitions for each value of ρ. The table I summarizes the mean number of wrong associations and the corresponding standard deviations. The goal model was extracted correctly in all trials. During three runs with $\rho = 1$ an additional (wrong) goal post hypothesis was initialized which did not affect the goal model on the higher level. In each run 1072 actual percepts were processed in 1285 frames. The number of wrong associations rises with higher ρ. However, even with ρ it accounts to less than 2% of all processed percepts.

In summary, the MHGM was able to handle dense as well as sparse false detections as expected. Dense false perceptions result in separate hypothesis and can be easily filtered out during the model extraction phase. Except for an insignificant amount, sparse false positives are filtered out by the percept buffer and dropped after a short period of time because no subsequent observations provide confirmation.

B. Approaching Goal

The aim of this experiment is to test the MHGM tracking performance in a dynamic game situation. In the initial situation of the experiment the robot is located between own goal posts. It is looking at a ball placed straight in the opponent goal and is therefore able to observe both goal posts. During

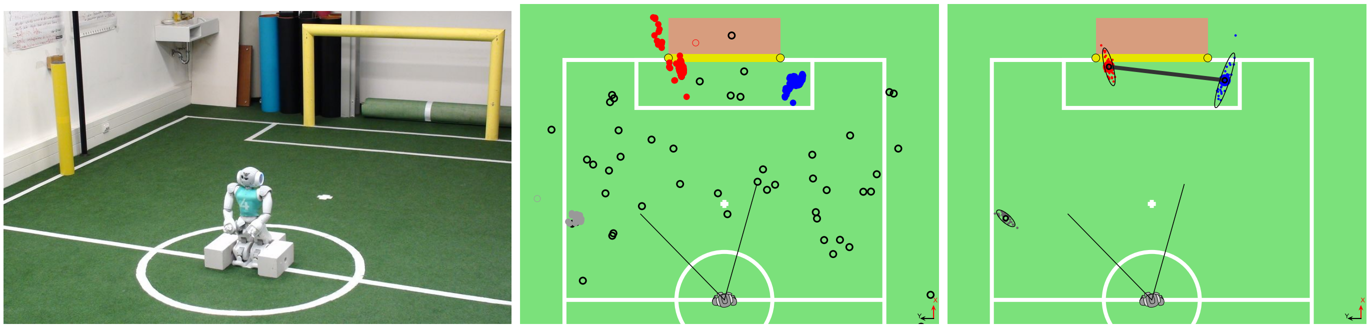

Fig. 3. The left Figure illustrates the experiment setup. The robot faces the goal and an additional goal post is placed to its right side. From the object recognition perspective, this post is identically to the *real* goal posts. The Figure in the center visualizes all percepts collected during the course of the experiment. The full circles illustrate perceived goal posts, whereby their color indicates the classification by the MHGM: red - left post, blue - right post, gray - unknown post, black - none (percept buffer). The circles with holes stand for artificially generated sparse false positive perceptions. The right Figure illustrates a snapshot of the state modeled by the MHGM at the end of the experiment. Drawn are the particle filter representing the goal posts with corresponding deviations as well as the extracted goal model. Similar to the Figure in the center, the colors of the particles indicate the classification of the hypotheses.

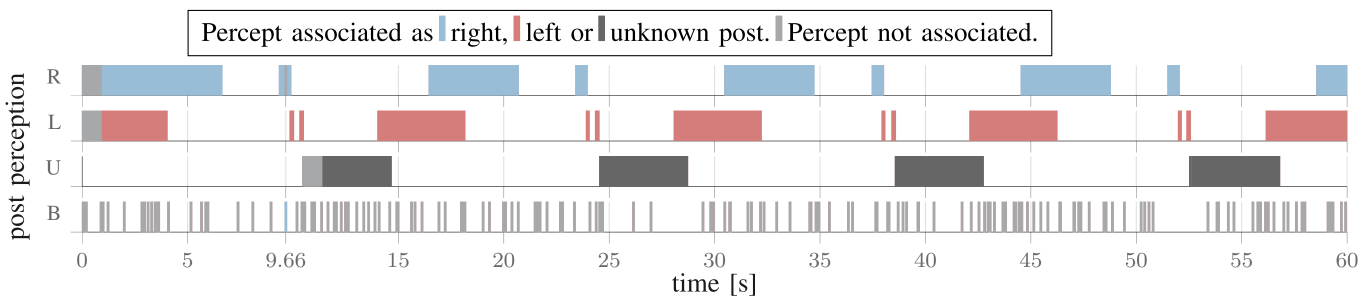

Fig. 4. This Figure illustrates the assignment of perceived goal percepts by the MHGM over time. The particular percepts are represented by colored bars and are sorted in four rows based on their actual origin (ground truth). The row **R** contains all percept triggered by the right goal posts, **L** by the left, **U** by the unknown third post and **B** contains the artificially generated sparse false positives which should be assigned to the buffer. The color visualizes the assignment of a percept by the MHGM. Blue stands for right goal post, red – left, black – additional unknown post, gray – no association (stays in buffer)

the experiment the task of the robot is to approach the ball and to keep track of the goal model. While approaching the ball the robot mainly is looking at the ball and periodically turning the head to collect more information. During the experiment, a tracking system is used to collect ground truth data about the robot position.

Based on the goal model it is possible to estimate the position of the robot on the field by simple triangulation. This naive method is applied to the goal model provided by MHGM to track the position of the robot during the experiment. The experiment was executed 10 times on the recorded perceptions. The Figure 5 illustrates the path of the robot on the soccer field. The path recorded by the tracking system is compared to the mean of the paths calculated based on MHGM together with the corresponding standard deviation.

The bottom graph of Figure 6 visualizes the percept association. It is worth noting that in all runs of the experiment no false assignations occurred. The two upper graphs show angle and distance to the left goal post. Thereby, perceptions and ground truth data are compared with the MHGM estimation. The goal post hypothesis based on a particle filter is easily following the goal posts perceptions despite interrupted detections and motion of the robot and noise is reduced significantly. The third graph from the top in Figure 6 shows the length of the main axes of the particle distribution modeling the left goal post in MHGM. With decreasing distance to the goal posts, the main axes decrease as well since the triangular error decrease. From second 32 no posts are perceived anymore and the axes increase again.

V. Conclusion and Future Work

In this paper we discussed the origins and components of perceptual uncertainties which are characteristic for humanoid robots and presented an approach for object modeling with a special emphasis on robustness regarding these uncertainties. The key idea is the separation of uncertainty in basic components like *noise*, *false detections* and *ambiguity*. Another aspect is the shape of uncertainty. This might be much simpler in local robot as in global coordinates. Based on these ideas, a multi-hypothesis tracking filter with delayed association is used to deal with noise and false positives, while the ambiguity is resolved based on the modeled hypotheses on the higher level.

This approach was applied to model a soccer goal (MHGM), which was tested within the RoboCup context. Presented experiments show that MHGM is able to handle uncertainty in a very effective manner and appears to be robust regarding massive noise and false positives. Our experiments were conducted in controlled scenarios and it still remains

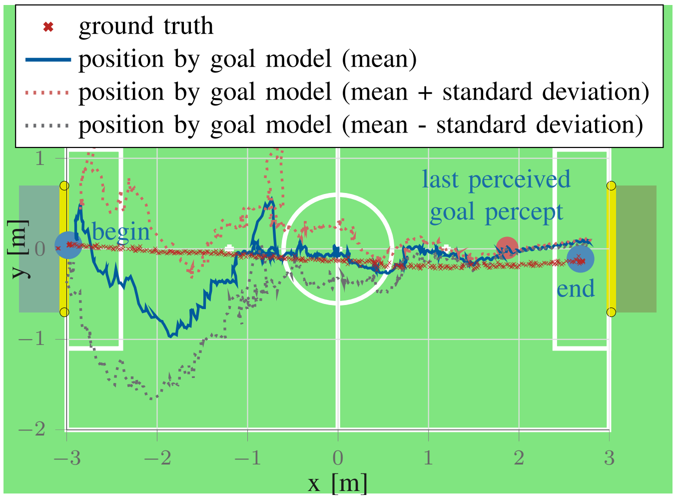

Fig. 5. The robot starts between own goal posts and walks to a ball placed in the opposing goal. The red line shows the ground truth of the robot position. Start and end points are marked accordingly. The blue graph shows the position of the robot calculated by a simple triangulation based on the MHGM and averaged over 10 runs. The dotted lines illustrate the corresponding standard derivation of the robot position. The point of the last perception of a goal post is also marked.

to be shown whether those results hold up in a real game situation. Nevertheless, all evidences so far indicate that our approach might be a step towards a solution which is robust regarding perceptual uncertainties and can relax dependencies on assumptions regarding error free vision and well calibrated robot kinematics. Although, MHGM can be used as a stand alone solution to track the soccer goal, it is planed to extend the model to the full situation including static and dynamic features, e.g., lines and ball.

References

[1] T. Kastner, T. Röfer, and T. Laue, "Automatic robot calibration for the nao," in *RoboCup 2014: Robot Soccer World Cup XVIII*. Springer, 2015, to appear.

[2] S. Tasse, M. Hofmann, and O. Urbann, "On sensor model design choices for humanoid robot localization," in *RoboCup 2012: Robot Soccer World Cup XVI*. Springer, 2013, pp. 380–390.

[3] T. Röfer and M. Jüngel, "Vision-based fast and reactive monte-carlo localization," in *Proceedings of the 2003 IEEE International Conference on Robotics and Automation (ICRA)*, D. Polani, A. Bonarini, B. Browning, and K. Yoshida, Eds. IEEE, 2003, pp. 856–861.

[4] M. J. Quinlan and R. H. Middleton, "Multiple model kalman filters: a localization technique for robocup soccer," in *RoboCup 2009*, J. Baltes, M. G. Lagoudakis, T. Naruse, and S. S. Ghidary, Eds. Berlin, Heidelberg: Springer, 2010, pp. 276–287.

[5] S. Tasse, M. Hofmann, and O. Urbann, "Slam in the dynamic context of robot soccer games," in *RoboCup 2012: Robot Soccer World Cup XVI*, ser. Lecture Notes in Computer Science, X. Chen, P. Stone, L. E. Sucar, and T. V. der Zan, Eds. Springer Berlin / Heidelberg, 2013, pp. 368–379.

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

[7] P. Guerrero and J. R. del Solar, "Improving robot self-localization using landmarks' poses tracking and odometry error estimation," in *RoboCup 2007: Robot Soccer World Cup XI*, ser. Lecture Notes in Computer Science, U. Visser, F. Ribeiro, T. Ohashi, and F. Dellaert, Eds., vol. 5001. Springer, 2007, pp. 148–158.

[8] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. The MIT Press, 2006.

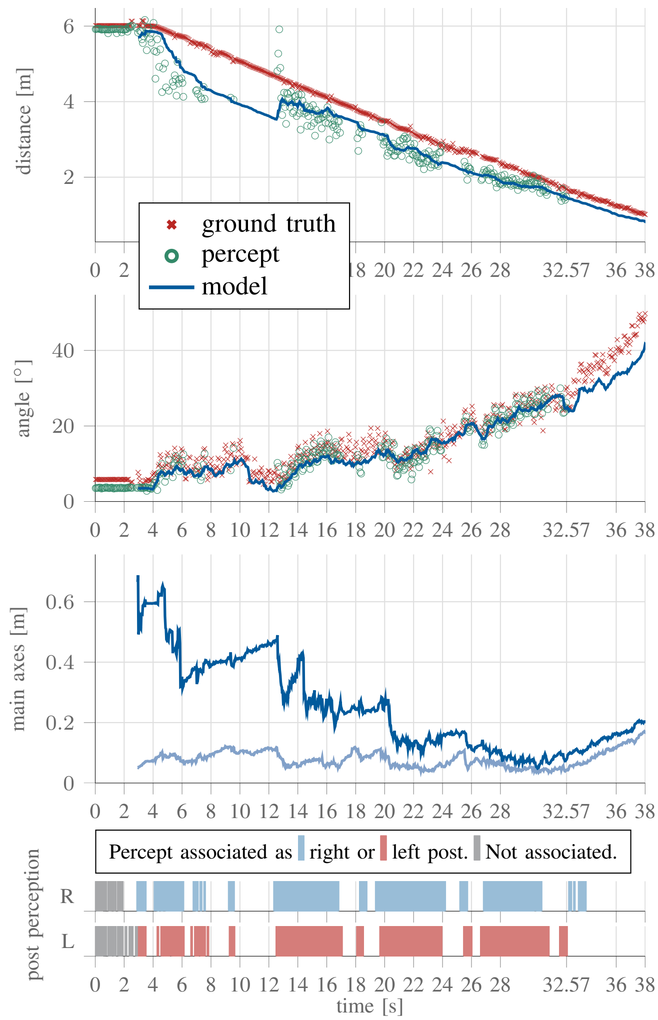

Fig. 6. The upper two graphs show the distance and angle to the left goal post. The MHGM consists of two hypotheses of different goal posts. Each hypotheses is represented by the mean of a particle filter. The upper third graph illustrates the main axes of the particle distribution the left goal post (dark – first principal component, bright – second principal component). The lower graph displays percepts of the left and right goal post. The colors illustrate their assignment to the MHGM (red – left, blue – right, gray – unknown).