Tracking for Following and Passing Persons *

Elin A. Topp and Henrik I. Christensen
Centre for Autonomous Systems
Royal Institute of Technology (KTH)
SE - 100 44 Stockholm, Sweden
{*topp,hic*}*@nada.kth.se*

***Abstract*— This paper presents a multiple target tracking approach for following and passing persons in the context of human-robot interaction. The general purpose for the approach is the use in Human Augmented Mapping. This concept is presented and it is described how navigation and person following are subsumed under it. Results from experiments under test conditions and from data collected during a user study are also provided.**

***Index Terms*— Human-Robot Interaction, Tracking, Following, Mapping**

I. Introduction

A key competence of a service robot is the ability to navigate in its human inhabited environment. This requires a map of the environment, that at some stage has to be acquired. At the same time there is a need to build a joint representation of the space, so that the human user can refer to locations and objects in the environment. Such references are parts of commands given to the system. Consequently, a metric map used for localization and navigation purposes has to be augmented with symbolic labels that can be used in future task specification dialogues. As we assume the user has no background in robotics, the mapping and labelling process has to be handled as naturally as possible for the user. A natural model is assumed to be a scenario in which the user guides the robot through the environment and specifies locations and objects as the map is acquired. Thus, a following functionality is required, which needs the user's trajectory. As we assume that in a real world scenario a number of persons might be around, the user has to be distinguished from other persons. Those other persons can in fact be seen as "human obstacles", and should be passed in an appropriate way. In this paper we present our approach to laser based tracking of multiple targets in the context of following one particular person through a populated environment.

A. Motivation

We describe an implementation of a tracking and following approach that is seen in the context of interactive mapping, or as we will call it "Human Augmented Mapping". With this term we subsume different aspects of interaction, tracking, following, concept learning and mapping.

*The work described in this paper was conducted within the EU Integrated Project COGNIRON ("The Cognitive Robot Companion") and was funded by the European Commission Division FP6-IST Future and Emerging Technologies under Contract FP6-002020.

We assume that a service robot is taken on a tour in an unknown environment and taught locations, rooms, and objects by its user. During this tour the robot builds up a map of the environment. Whenever the user specifies a new location, room or objects, a respective label is created and assigned to the map. Later these labels will get semantic connotations, so that reasoning about the environment becomes possible. Thus, the term "Human Augmented Mapping" represents an integration of automated map generation (simultaneous localisation and mapping, SLAM), and learning of environmental concepts.

One key functionality in this context is following a person around within an interaction context. As we assume further, that the service robot is navigating in a populated environment such as a house or office building, an approach to passing a person, or "socially acceptable navigation" is needed. With this term we mean the ability to differentiate between obstacles and persons so that an appropriate behaviour can be selected. An example as introduced by Hüttenrauch et al. [5], would be a bystander meeting the person with the robot following. In such a situation one would want the robot to act "socially" and keep a certain distance, not only to the person it is following, but also to the bystander. On the other hand the robot could navigate in close proximity to objects. To achieve such a behaviour the system needs to keep track of all persons in a certain area of interest, and not only the user. Therefore an approach for tracking multiple targets from a moving robot is needed as the basis for the following behaviour. As laser range data cover a wide angle, are easy to handle, and calculations on them are fast, we see them as the most useful basic sensory data for this purpose.

B. Related Work

There is an abundance of literature on people tracking, in particular using laser range finders and colour based vision. The majority of work in robotics relies on the use of SICK laser scanners. Most work concentrates on either tracking one person – a user – for interaction, as presented by Kleinehagenbrock et al. [6], or multiple targets, but without the intention of interacting [10]. Kleinehagenbrock et al. [6] use a laser range finder to detect a user and verify this hypothesis with the help of vision based face detection. The two cues are integrated with a multi-modal anchoring technique.

Few publications concentrate on the relationship between user and robot in motion in a dynamic environment.

One paper that discusses this, is presented by Prassler et al. [9]. The authors coordinate the movements of a robotic wheelchair with an accompanying person in a busy large scale environment. Other people, i.e. bystanders, are seen as dynamic obstacles that reduce the free space for navigation. In contrast to that, the present paper focuses on a rather narrow space, as given in an office environment or a regular house, where crowds are seldom, but a few bystanders besides the user might be present.

C. Overview

The paper is organised as follows: In section II we explain our approach to tracking multiple targets in the context of following and passing persons. Section III gives an overview of hardware and our system architecture. Section IV presents the results from tests of the system under experimental conditions, and from test runs on data collected during a user study conducted at our laboratory in September 2004. In section V we draw conclusions and present some ideas for future improvements.

II. Tracking for Following and Navigating

This section deals with the criteria that a tracking system needs to fulfil to serve our purposes. We will also give some ideas on why we think that a multiple target tracker is a useful approach to solve both problems which, as mentioned before, are

a) tracking a user in a natural environment to allow following, and
b) keeping track of persons in the vicinity for socially acceptable navigation.

These two purposes lead to slightly different criteria for quality measures of the tracking approach. Two different types of tracker failures can occur with respective effects on the reliability of the system. These two types are the loss of a target and a confusion of different targets due to ambiguous data association. We discuss those failures with respect to their effects.

A. Tracking challenges

In case of a following scenario the output of the tracker needs to be analysed to find the one and only person to follow. In this case the purpose of the multiple target tracking is to find all user hypotheses and later distinguish the user from other persons.

One possible tracker failure would be to lose the target associated with the user due to detection failures over a certain period of time. If in such a case the target is removed and after a while replaced by a new one, the tracker output can still be used to define this new target as the user, as the target is probably still in the area where a user is expected. Otherwise a complete target loss could lead to an error state and thus be handled appropriately.

More critical for the following scenario is the confusion of the user target with another one, possibly another person moving in a different direction. This is a situation that is to be avoided by any means, as the system would not detect any error and would start following the wrong person.

For the purpose of passing persons the criteria are slightly different. Here it is not important to know which of the persons is associated with which target, but targets must not be lost when in fact the respective person is still around. In such a situation a person would appear as an arbitrary obstacle to a general obstacle avoidance routine.

Considering those aspects, a robust tracker that allows to distinguish between targets is a solution to both problems, as they could occur at the same time. This would be the case when the system is following one person around while reacting appropriately to the presence of other persons.

B. The tracking system

Our laser tracking system is based on the approach presented by Schulz et al. [10]. Their system uses leg detection and occupancy grids to detect people and distinguish them from other objects by movement. Detected features are associated to tracked targets with a sample based joint probabilistic data association filter (SJPDAF). Using this they can track multiple targets from a mobile robot, while motion compensation is done by scan matching. We adopted the idea of using the SJPDAF approach for tracking and associating, but in contrast to Schulz et al. our detection and tracking allows handling of people standing still, which is useful for interaction.

C. Detecting humans in laser data

A common method to detect humans in laser data is to look for leg hypotheses, as done by Feyrer and Zell [3], Kleinehagenbrock et al. [6] and Schulz et al. [10]. The laser range data are analysed for leg sized convex patterns, either one of them or two at a reasonable distance from each other. Other systems rely on body shape as presented by Kluge [7], or in our previous work [12]. In this case a single "person sized" convex pattern is extracted from the data as a person hypothesis. The choice between the two approaches is often determined by the height the used laser range finder is mounted at. We think that accepting leg patterns only is a rather strong constraint, as in this case a person wearing a skirt or baggy trousers would not be classified as person. Therefore we allow three types of patterns. These patterns can be classified as *single leg, (SL)*, *two legs appropriately separated, (TL)* and *person-wide blob, (PW)*. As accepting these patterns all the time would potentially generate a large number of false alarms, a rule based approach is adopted for the generation of new person hypotheses,

- TL and PW are accepted as features at any time they occur,
- SL are only accepted when they are close to an already detected and tracked target.

The latter constraint is based on the observation that a single leg pattern can only be seen for a short period of time when the leg of a moving person occludes the other. Therefore all other SL patterns are ignored, as they are unlikely to belong to a person. On the other hand we need to allow the SL pattern for a smooth tracking of the targets that have been already accepted.

D. Tracking and associating features

As mentioned above we use SJPDAFs to associate targets and features in a probabilistic framework. Each feature $z_j \in \{z_0, z_1, z_2, \ldots z_n\}$ is assigned a posteriori probability β_{ij} that it was caused by the target $x_j \in \{x_1, x_2, \ldots, x_m\}$. The feature z_0 represents the case that a target was not detected at all. The computation of the β_{ij} is based on a sample representation for the targets. Each target x_i has its own sample set for state prediction and is updated according to β_{ij}.

The sample space is composed of the state (x,y,v,θ) of their respective target, where (x,y) refers to the position, v is the translational velocity and θ the orientation relative to the robot. A first order Taylor expansion is used for the motion estimation. Although the data association method is meant to handle a number of objects it originally assumes a fixed set of targets. As we can expect that only very few new targets would enter or leave the scenery at exactly the same time, the method is still a valid way of solving the association problem.

Detection misses are handled depending on their duration. If a target is occluded by another one and is therefore not detected, it remains in the list of targets and its position is continuously estimated for a period of time. After this time period we assume that it has disappeared. If in this context a target is removed because of an occlusion, and is detected again afterwards, it is assigned a new target identifier. We assume here, that a user is cooperative and does not deliberately hide behind another person or object, when being followed by the robot.

In order to handle motion compensation we currently use the scan matching technique presented by Lu et al. [8], which gets odometer information as an initial estimation. A good position estimation for the moving robot is crucial for the data association, especially in cases of occlusions. If the position of a target cannot be updated by from measurement data, the particles for this target predict the new state according to their prediction rules in combination with the estimated movement of the robot. Especially in the case of a rotation, a large error in the estimation can cause ambiguities in the data association. This can be explained by a rotational shift, that would be erroneously applied to all target position predictions, and new target hypotheses appearing due to an actual rotation.

E. Interpreting the tracker results

As our main interest is not the tracking approach itself, but its performance in the context of person following – or motion coordination – and person passing, the tracker results need to be interpreted accordingly. To keep the tracking approach as independent as possible, the tracker itself does not know about this interpretation. All it produces is a set of targets in each step. This set of targets is currently fed into what we call the person handler. This person handler updates the position information in its person set and assigns each target/person a certain state, comparable to the user states we defined in previous work [12]. Those states can be described as follows:

- PERSON: This is the initial state, as all features that match our pattern classification are assumed to belong to a potential person of interest.
- MOVING: As we do not have other means of classification for the user right now, we need to have some more information. Thus, we set the state for a person to "moving", whenever a certain distance was covered by the target.
- USER: To determine the user we currently use a very simple rule: The closest moving person within a certain distance and angular area relative to the robot is assigned the "user" flag. Only one user at a time can be present and once a person target gets the user flag, it will keep it, until it disappears from the scene.
- GONE: A target that has been removed from the set of targets is set to "gone" in the person handler. This allows higher level processing of this state, for example, producing an error message when the user target is lost.

The criteria for picking the user are currently designed to allow for easy tests that do not rely on other modalities for confirmation such as spoken dialogue.

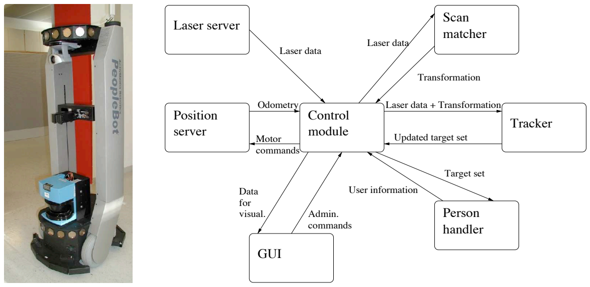
F. Following

Our first approach for tests of the following behaviour is a distance based control loop. The desired speed and turn rate are computed depending on the distance to the user and modified by the distance to obstacles. This results in a careful manoeuvring in narrow environments and higher speeds in open spaces. This simple approach could be replaced by more sophisticated control methods that are based on the idea of a desired position relative to the user, as for example used by Prassler et al. [9]. However, for the tests of the tracking system our straight forward approach seems appropriate.

III. SYSTEM SETUP

Our testbed is an ActivMedia Performance PeopleBot, shown in figure 1a). The robot is equipped with a SICK LMS-200 laser ranger finder mounted at a height of about 33cm. This type of range finder can provide data transmission rates up to 38400 baud over a regular serial connection. Using a respective Serial/USB converter we can obtain data at the maximum speed of 500000 baud. In the latter case we pay for the high data rate with an unstable communication, which is a known problem[1]. Therefore we run the tracker with different transmission rate configurations in our experiments. Other available sensors such as the camera, sonar or infrared sensors are currently not in use. For communication with sensors and the robot controller we use the Player/Stage[2] robot control software libraries.

[1] http://www-robotics.cs.umass.edu/segway/sick.html
[2] http://playerstage.sourceforge.net



a) b)

Fig. 1. a) Our ActivMedia Performance Peoplebot "Minnie", equipped with laser range finder. b) Architecture of our track and follow system.

Software architecture

In Fig. 1b) our control architecture for the tracking and following behaviour is described. The central part is a control module that handles incoming sensor information and sends it to the respective modules for interpretation. Resulting commands for the motors are generated and sent to the robot controller via the interfaces of the Player robot control server. Other important modules are the tracking module and the person handler. The tracker bundles the detection of new features and the actual tracking. Its output is the updated set of targets, which is then dealt with in the person handler. The latter is responsible for labelling the targets according to their level of interest for the interaction. The scan matching for motion compensation is provided by a separate module to make an easy exchange possible. This allows us also to substitute the scan matcher by pure odometer information for tests and comparisons.

For visualisation, diagnostics and comfortable handling a graphical user interface is connected to the control module.

IV. Experimental results

To test the tracking approach in the discussed context, we used three different scenarios. One tests following and tracking of multiple targets in an artificially emptied "room". The other two reflect the behaviour of the tracker in a "real world" context. As we observed differences in the quality of the results, we present the two test types separately, referred to as setup #1, #2, and #3 respectively. The tests in "real world environment" concentrate more on the following aspect, as the passing approach itself is subject to current work and could therefore not be used as extensively in experiments.

A. Experimental setup #1

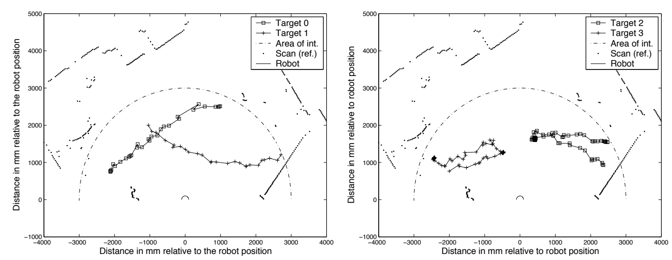
In order to make sure that the number of persons present was controllable at any time during experiments, we defined an empty area by setting up a number of large plywood planks and cardboard pieces as "walls" for the experiments that involved a moving robot. We then defined a number of test scenarios as follows:

- robot not moving, one person present,
- robot not moving, two persons present, occluding each other,
- robot moving independently, up to three persons present, and
- robot following one person.

With these tests we aimed to test the tracker under different test conditions . Regarding our quality measurements presented in section II-A we were mainly interested in problematic situations that might lead to confusions or the loss of a target. We therefore asked our test persons to walk at different speeds, cross each other in the field of view of the robot on purpose, "meet" in the middle of the room, "chat" and separate again, or perform unexpected changes in their moving direction.

The laser range finder was set to a data transmission rate of 38400 baud to guarantee stable transmission and to determine, if this speed was enough for our purposes.

During the tests all occlusions were handled correctly and no target was lost. This result could be confirmed by different tests under similar circumstances with the same models for movement and state prediction.

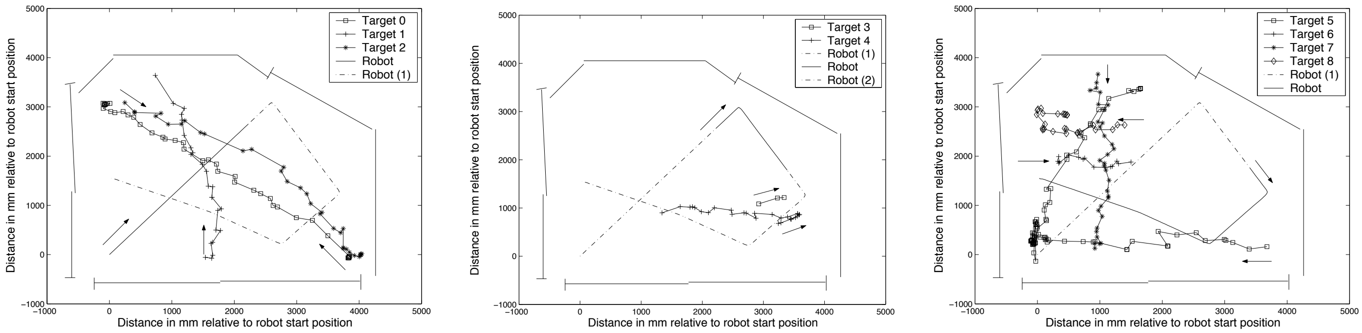


a) b)

Fig. 2. The trajectories for a test with two persons moving in the field of interest, with the robot standing still. The dashed line marks the area of interest, the small half circle at position (0,0) represents the robot. The dots show a reference scan of the environment. a) The two persons cross the area of interest, with target 0 being occluded by target 1. The trajectory for target 1 starts rather late after walking into the area. This can be explained by the accepting rule for features. Taking a look at the data at the corresponding time steps shows that the person in question was represented by a single leg pattern for a couple of steps. b) The two persons walk into the middle of the area, stop at a comfortable "chatting" distance (about 80cm) and separate again.

1) Robot still, one person: In this test scenario one person crossed the field of interest (in this case the area described by the laser base line (y=0) and a radial distance of three meters) nine times, at varying speeds. The target was not lost at any time. It was always classified as moving person and got the user flag when entering the area where a user would be expected.

2) Robot still, two persons: Two persons crossed each other in front of the standing robot, went out of the area of interest and came back. They met in front of the robot, "chatted" and separated. Figure 2 shows the resulting trajectories. Again, the area of interest was set with a radial distance of three meters. In this case, the surrounding environment was the natural lab environment, but it was made sure that no disturbing objects were in the field of interest. This was possible, because the robot did not move. This test gives an



a) b) c)

Fig. 3. Person and robot trajectories in the x/y-plane of an experiment with up to three persons in the area at the same time and the robot moving independently. The resulting trajectories are presented in three images for clarity of presentation. The robot trajectory is shown as a solid line for the time period representing the events in the area. The rest of the robot's trajectory is shown dashed for convenience. a) Three persons cross the room in different directions. The "long steps" in the trajectories (in the starting steps for target 0, in the middle part of target 2's trajectory and in the last steps of target 1) occur due to occlusions. b) When the robot turns in the upper corner of its path, it loses the recently detected target 3 out of the field of view. Target 4 performs an indecisive behaviour by turning around and going back after a few steps. c) Target 5 remains in the scene for almost the whole time period shown in this graph, crossing the area from right to left, standing for a while in the bottom left corner and then continuing "up", being occluded by target 7 for a short moment.

example of the tracker being able to handle the short term occlusion of two persons passing each other. Both targets are classified as moving persons and target 1 gets the user flag, when entering the respective zone in front of the robot. The trajectory for target 0 seems to stop clearly within the area of interest, as the person gets occluded by some object indicated by the respective scan data points in the image. As the person does not come out of this hiding place for a while the system assumes the target is gone. As well for the "chatting" scenario the tracker could handle the situation, which shows, that if two targets get close to each other, but are clearly distinguishable, no confusions occur. Again, one of the targets (target 2) gets the user flag as it enters the respective zone first.

3) Robot moving, three persons: This test is the most relevant for our purposes as it shows the abilities of the tracker running on the moving robot, together with the target classification that would make the robot follow one of the persons. Figure 3 shows the resulting trajectories from one of the tests covering this type of scenario, with three persons moving around while the robot is crossing the area. For this particular test we set the area of interest to a radial distance of eight meters. This means, that the whole "room" is in the field of interest. The robot moved straight across the area until it detected one of the walls at a certain distance. Then it turned randomly to the left and to the right until it had enough free space in front to continue. In the first part of the scenario target 1 gets the user flag. This happens due to its position closest to the robot when it enters the "user zone". For the second part of the test target 3 did not remain visible long enough to be classified as a moving person, but target 4 gets the moving person and user classifications. For the last part target 5 is found as user and keeps the flag while it is present. The targets 6 and 7 are classified as moving persons, but not as user, as target 5 is still around. When target 5 steps out of the field of view, the user gets lost, but immediately afterwards the newly arrived target 8 gets the user flag.

4) Robot following one person: To show the tracker's ability in a following scenario, the system was set in the respective mode and followed one person for about three minutes. During this time period the user changed her walking behaviour (speed and direction) frequently, sometimes came very close to the robot, so that it had to move backward, and stepped close to the walls of the empty room used in this experiment again. This test over a period of three minutes shows, that our motion model is able to handle arbitrary movements quite well, as the user was not lost at any time.

From these tests we can conclude, that under test conditions our approach can handle the situations we are interested in. Nevertheless, running the tracker with slight changes in the motion model on the same data sets for a number of times showed that there are situations in which the tracker fails, due to a seriously wrong prediction of the further movement direction of a target in combination with a detection miss for the same target. This indicated that it might be useful to switch to a more sophisticated motion prediction model as derived by Bennewitz et al. [1] or Bruce and Gordon [2]. For the moment though it seems more appropriate to look into the problem of the detection itself, as the following results indicate.

B. Experimental setup #2

As we cannot assume such test conditions all the time, we tested the tracker on data collected during a comprehensive user study conducted at our laboratory. The user study was a Wizard-of-Oz experiment and is described in detail by Green et al. [4]. One important fact to note about this kind of experiment is that the robot was actually controlled remotely, while the test subject was told that the system performed autonomously. The scenario for the experiment was a guided tour through a living room. Subjects got the task to ask the robot to follow, present different locations and objects in the room and test the robot's understanding by sending

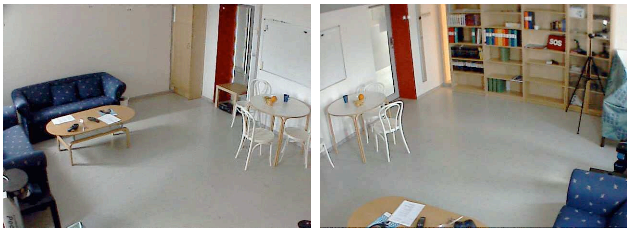


Fig. 4. The experiment environment ("living room") seen from different perspectives

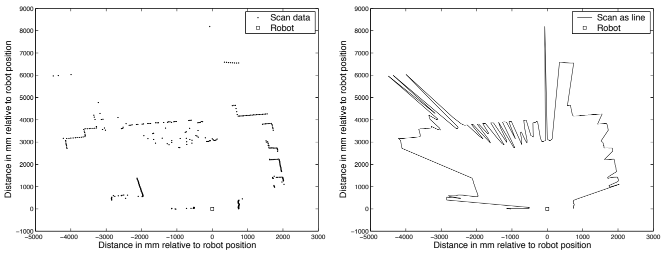


Fig. 5. The raw laser data (top) and the same data represented as polyline to show the data points in their angular order. The two peaks right in front of the robot are caused by the subject's legs, while the other peaks result from the table and chairs, that belonged to the experimental scenario.

it to learnt locations and objects. The study includes data from 22 experiments. Laser range data was collected in all experiments at a data transmission rate of 500000 baud. Due to the mentioned communication stability problem (see III), not all of the experiments could be recorded completely. Still, we have a body of a couple of hours of experiment sequences, since every experiment lasted between 10 and 20 minutes. Figure 5 shows a raw scan taken from a typical start position during the tests.

Running the tracking system on the data from the experiments showed, that performance in this kind of real world environment was significantly worse than expected after the results from the previously reported tests. The user target got confused with other targets rather frequently, which we defined as a critical failure in following.

Taking a closer look at the process revealed the reason for the confusions. This is best explained with an example, as all those confusions occurred for exactly this reason.

As stated in section II we allow static targets, as we assume that this is reasonable in an interaction context. In fact, the experiments with the "inexperienced users" confirmed this assumption, as many of the subjects stood still for quite a while (up to 50 seconds) repeatedly.

The images in figure 5 show a clear resemblance between some of the patterns and the subject's legs, even if some of them appear too pointy. Still, such patterns can fall under the classification thresholds for legs. As we cannot assume a completely smooth representation for the target's movement (as this would conflict with the Sampling Theorem [11] and the laser ranger finder's angular resolution), differing pattern parts can get associated to leg pairs, when the robot moves

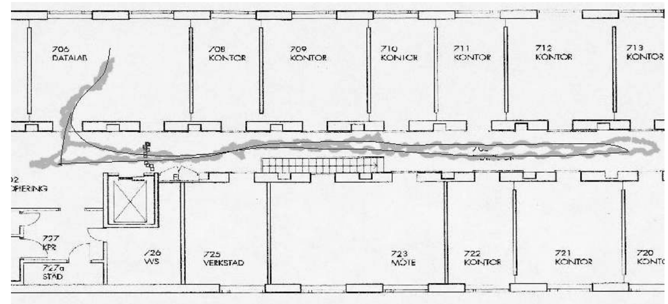


Fig. 6. The trajectories of the user (grey, thick line), the following robot (black, thin line) and a bystander crossing the way of the robot (small squares) between the robot and user.

and changes perspective. Thus, the erroneously detected target(s) start to "move". The respective sample set picks up the motion estimation and predicts a new position. If the robot's viewpoint changes such that the "target" is not detected for a while, the predicted state gets more and more ambiguous. With the particles spreading toward the actual position of the user target and the appearance of a new (erroneous) target, the statistical approach is likely to confuse the feature-target association. As a consequence of such a confusion, the tracker needs a few steps to recover, i.e. retrieve its certainty, which is even more difficult when the user stays close to distracting objects.

As the task for the subjects was to show the robot around in a furnished room, it is scenario immanent that the user moves around between objects in the room. On the other hand, we could state, that in situations where the user was clearly distinguishable from disturbing objects, and those disturbing objects were detected reliably, the tracker and data association performed as expected. Occlusions were also handled properly in these situations.

C. Setup #3: Following through the office building

With this test we showed, that the system is still suitable for "real world" conditions, if the disturbances can be reduced to a minimum by the choice of environment. The robot followed the test person out of the laboratory and along the hallway, covering a distance of about 25 meters, and returned – still following – to the laboratory. On the way back a bystander was asked to cross the way between the robot and the user. Figure 6 shows the part of the office building together with the trajectories. The experiment took approximately four minutes and a distance of about 50 meters was covered, including two door passages. A total number of 26 targets was detected throughout the whole time period, one was accidentally classified as "moving", but did not get confused with the user. The user target was tracked reliably over the complete time period and one occlusion of the user by a crossing bystander was handled as expected. The bystander target was classified as moving person correctly, so a respective person passing method could have handled the situation appropriately. In our test case the robot slowed down a bit, due to the influence of obstacles on the speed.

Summarising these tests on "real world" data we observed that

- the approach for tracking and data association is still a valid method for tracking multiple targets in the context of following a user or passing persons.
- the approach is highly sensitive to motion models, but the choice of a good motion model does not seem to be as critical as the reliable detection of actual targets.
- problematic situations occur in "real world" scenarios, i.e., cluttered environments, when vicissitudinous false alarms lead to confusions.

Judging from these observations, we assume that we can improve the system for following and passing persons significantly by introducing other means for the detection of targets. Within the context of "showing the robot around" we have to deal with an unknown, cluttered environment. From preliminary analysis of the user study we can allege that persons in this context move differently compared to results from observations in long term experiments on a larger scale. Subjects tended to move to a certain location, stop and move around in a small area, to "explain" things to the robot. This type of movement seems rather stochastic, compared to the motion models that hold for long distance movements. Therefore we think that improving the detection to eliminate confusing false alarms is a better way to improve the system for our purposes.

V. Conclusion and future work

In this paper we presented our approach for tracking multiple targets with a mobile robot in the context of following a user and navigation in a human populated, but not too busy, environment as office buildings or regular houses. In this context we introduced the term *Human Augmented Mapping* with which we subsume aspects of various disciplines such as human robot interaction, people tracking and following, localization and mapping. We described experimental results for the tracking approach under test conditions in an environment that did not contain any disturbing objects. Further we described the results from tests on data collected during a user study in a typical scenario and presented an experiment with our robot following a user around our laboratory for a certain time period.

As a general outcome we can state that our method is capable of handling a typical scenario. The multiple target tracking approach allows differentiation between the user that has to be followed, and bystanders who are not immediately involved in the interaction of user and robot, but still need to be detected and included in navigation planning.

A couple of confusing situations occurred during our tests in the "real world". The analysis of these situations revealed a need for an improvement of the detection of people in cluttered environments. There are different options to realise such an improvement. One is to base the detection on multi-modal sensory information, for instance by adding vision-based face detection for hypothesis verification. Another one is to start a detailed analysis on patterns caused by persons in laser data. We will work on such an analysis as a supplementary subject in the near future.

In general it might be helpful to investigate the effect of a context dependent tracking model that include knowledge about the environment, which requires map information. This and the fact that our approach is still valid for the purpose of person following indicate that it is useful to continue with the next steps toward a Human Augmented Mapping system and to work on improvements of the tracking system iteratively. These next steps involve the integration of a mapping method together with some interactive means of labelling the achieved map. This is the main subject of our current work.

Acknowledgements

We would like to thank H. Hüttenrauch, A. Green, and K. Severinson-Eklundh who were responsible for the user study as well as E. Pacchierotti for some helpful discussions on socially acceptable navigation.

References

[1] M. Bennewitz, W. Burgard, and S. Thrun, "Using EM to learn motion behaviors of persons with mobile robots", in Proc.of the Conference on Intelligent Robots and Systems (IROS), Lausanne, Switzerland, 2002.

[2] A. Bruce, and G. Gordon, "Better motion prediction for people tracking", in Proc. of the International Conference on Robotics and Automation (ICRA), New Orleans, USA, April 2004.

[3] S. Feyrer, and A. Zell, "Robust Real-Time Pursuit of Persons with a Mobile Robot Using Multisensor Fusion", in Proc. of the 6th International Conference on Autonomous Systems, 2000.

[4] A. Green, H. Hüttenrauch, K. Severinson-Eklundh, "Applying the Wizard-of-Oz Framework to Cooperative Service Discovery and Configuration", in Proc. of the 13th IEEE International Workshop on Robot and HumanInteractive Communication, September 2004.

[5] H. Hüttenrauch, and K. Severinson-Eklundh, "Fetch-and-carry with CERO: Observations from a long term user study with a service robot", in Proc. of the 11th IEEE International Workshop on Robot and Human Interactive Communication, September 2002.

[6] M. Kleinehagenbrock, S. Lang, J. Fritsch, F. Lömker, G.A. Fink, and G. Sagerer, "Person Tracking with a Mobile Robot based on Multi-Modal Anchoring", in Proc. of the IEEE International Workshop on Robot and Human Interactive Communication (ROMAN), Berlin, Germany, September 2002.

[7] B. Kluge, "Tracking Multiple Moving Objects in Populated Environments", in Lecture Notes in Computer Science, volume 2238, eds. Hager, Christensen, Bunke, and Klein, pages 25–38, Springer Verlag, Heidelberg, October 2002.

[8] F. Lu and E. Milios, "Robot pose estimation in unknown environments by matching 2d range scans", in Proc. of the IEEE Computer Vision and Pattern Recognition Conference CVPR, 1994.

[9] E. Prassler, D. Bank, and B. Kluge, "Motion Coordination between a Human and a Mobile Robot", in Proc. of the 2002 IEEE/RSJ Intl. Conference on Intelligent Robots and Systems, Lausanne, Switzerland, October 2002.

[10] D. Schulz, W. Burgard, D. Fox, and A.B. Cremers, "Tracking multiple moving targets with a mobile robot using particle filters and statistical data association", in Proc. of IEEE International Conference on Robotics and Automation, Seoul, Korea, May 2001.

[11] C.E. Shannon, "A mathematical theory of communication", Bell System Technical Journal, vol. 27, pp. 379-423 and 623-656, (July and October, 1948)

[12] E.A. Topp, D. Kragic, P. Jensfelt, H.I. Christensen, "An Interactive Interface for Service Robots", in Proc. of the IEEE International Conference on Robotics and Automation, New Orleans, USA, April 2004.