Revision Summary

"Incorporating Perception Uncertainty in Human-Aware Navigation: A Comparative Study"

Zeynab Talebpour, Deepak Viswanathan, Rodrigo Ventura, Gwenn Englebienne and Alcherio Martinoli

May 27, 2016

The main issues pointed out by the reviewers have been addressed in this version of the paper and a numbered list of how each point has been resolved or explained can be found in the following.

1. Concerns about the tracking model:
   We have made appropriate revisions to equation 1 and 2 consistent with our implementation and the standard Bayes filter tracking model of [21].

2. Learning of $P(\mathbf{O}_t|\mathbf{X}_t)$ for the locations and configurations of the state space:
   We agree with the reviewers that the details of the observation function/likelihood function has not been explained in detail. A motivated reduction in the state representation is required to be able to learn the observation function from labeled data. We have built on the approach of [22] to generalize such an observation function over state features such as relative location of the camera and the people in the environment. The details of the method are outside the scope of the current paper.

3. Evaluation concerns regarding the ground truth:
   This is a completely valid concern and as pointed out in the future work of our paper, we are working on improving this absolutely important aspect of providing a precise and reliable ground truth for tracking moving people in real environments. However, with the marker-less tracking system that we had at the time we could not do any better.

   We were aiming to understand the overall behavior of the robot when using our proposed method and we have records of a modified behavior of the robot in terms of human-aware navigation in all of the scenarios, knowing the positions of the people (scenario 1 and 3) or their overall trajectory (scenario 2). In other words, although the position of the moving person is not precisely known at all times during all experiments, the different behaviors of the robot when adopting each (social) navigation method is very informative and we can conclude the effectiveness of our method based on this. Nonetheless, there is no question about the necessity of improving our evaluation methods.

For scenario 2, the modified trajectory of the robot shows how CHA can be a remedy to noisy perception by considering uncertainty. In Figure 4.b CHA has resulted in a trajectory that is smoother and deviates from the straight line, whereas, BN and DHA (due to delays in person tracking) make abrupt changes when encountered closely by the person.

4. Reporting sigma values:
   This was previously reported. However, we made it more clear by adding more information to the last paragraph of section IV (Human-aware Navigation Model).

5. Motion transition model is limited to static and low dynamic scenes:
   A better motion model will provide a more accurate estimation of the location of the people. Our current model computes a principled form of uncertainty in a multi-person environment. With the uniform motion model we have used, we have a less accurate estimate of the state in high dynamical scenarios compared to a more sophisticated model. Nevertheless, this is not a limitation for the current work. The focus of this work was to develop a principled human-aware navigation approach considering uncertainty in perception. From this perspective, the use of a simple motion model is not a constraint to our human-aware navigation model. But this is definitely an area we could explore more in future works.

6. Explanation of Figure 7*c* in terms of lower accumulated social cost for CHA:
   Considering all experiments of scenario 3 the average accumulated social cost associated to DHA is larger than that of the CHA (2923 for DHA vs. 2772 for CHA).

7. Title capitalization inconsistencies:
   All titles for sections and subsections follow the same format now.

8. Adding the trajectory of the moving human:
   Figure 4b is changed to indicate the trajectory of the moving person over time. However, we emphasize that this trajectory is not obtained by the tracker, but is given as a plan to the person involved in the experiment. This is done using static ground markers for start and end positions of the person's trajectory and the person is asked to move with a constant speed in a straight line. For the other 2 scenarios involving static persons, their ground truth positions were also known and indicated using ground markers. As stated in the paper, the ground truth position given by the tracker is not very precise for the case of the moving person due to delays and larger position errors. Improving this aspect is a vital future step in our research.

9. Why does BN keep a larger distance to the person depicted by the red circle in Figure 4, compared to the CHA method?
   This is due to the robot being ignorant of that person and trying to take the shortest path to the goal in BN. In other words, the robot does not differentiate between people and obstacles and tries to avoid people using

its local sensors without any social path planning. Therefore, in this case, the robot does not plan ahead for having a socially-accepted navigation behavior when encountering the person in blue and moves in a straight line instead. This results in a larger distance to person shown in green, while CHA assigns some social costs to areas around both humans which causes the robot to deviate its path from a straight line.

10. Lack of statistical analysis:
    This is an important point that will be considered in the future. However, since we had three scenarios and five methods we could not afford more tests with real robots in the limited time that we had. This gives a total of $5 \times 3 \times 5 = 45$ trials which were partially reported for the sake of concessions (we did not report clustering methods for scenario 2 and 3 since they were shown to be inferior to CHA in scenario 1).