# Spatially Unconstrained, Gesture-Based Human-Robot Interaction


Guillaume Doisy
Department of IEM
Ben-Gurion University of the Negev
Beer Sheva, Israel
Email: doisyg@post.bgu.ac.il

Aleksandar Jevtić
Robosoft
Bidart, France
Email: aleksandar.jevtic@robosoft.com

Saša Bodiroža
Institut für Informatik
Humboldt-Universität zu Berlin
Berlin, Germany
Email: bodiroza@informatik.hu-berlin.de



***Abstract*—For a human-robot interaction to take place, a robot needs to perceive humans. The space where a robot can perceive humans is restrained by the limitations of robot's sensors. These restrictions can be circumvented by the use of external sensors, like in intelligent environments; otherwise humans have to ensure that they can be perceived. With the robotic platform presented here, the roles are reversed and the robot autonomously ensures that the human is within the area perceived by the robot. This is achieved by a combination of hardware and algorithms capable of autonomously tracking the person, estimating their position and following them, while recognizing their gestures and navigating through environment.**


## I. Background

The space where human-robot interactions can take place is restricted by the limited perceptual capabilities of robot's sensors and/or the lack of mobility of the robot. By using external sensors and creating intelligent environments, such as in [1], robots are able to follow humans. Recent release of the affordable depth sensors lead to similar solutions that do not rely on external sensors [2]. Natural interaction techniques, such as gesture recognition, assume that the relative position of the sensor and the human remains unchanged during the interaction [3], [4]. Such techniques need to be improved to be used in less constrained scenarios [5]. This work presents a robotic platform benefiting from the latest development in person following and gesture recognition to expand the space where HRI can take place.

## II. Framework for Space Extended HRI

### A. Person Tracking

Person tracking is performed using the Kinect sensor and the Microsoft Kinect SDK. The limited $57°$ horizontal viewing angle of the sensor does not allow for person tracking $360°$ around the robot. To overcome this limitation a visual servoing control law was developed to rotate the pan-tilt mechanism (PTM) and the Kinect sensor in the direction of the tracked person. Experimental results showed that this visual control law ensures continuous person tracking when the person is moving around the robot. The latency of the Kinect sensor is too high to allow for continuous tracking when both the robot and the person are moving: in the case of fast movements the system is not reactive enough and loss of tracking can occur. This problem was solved by adding a term in the PTM control law which compensates the angular speed of the robot given by the odometry (i.e. making the PTM rotate with a speed opposite to the angular speed of the robot). For a formal description of the control law, please refer to [2].

### B. Person Position Estimation

It is possible to compute the absolute position of the tracked person by using a series of frame of reference transformations, if the position of the person in the Kinect's frame of reference, the rotation angle of the PTM and the absolute position of the robot estimated from odometry are known. In order to eliminate the noise due to the vibrations of the robot structure a $15cm$ jitter filter and a low-pass filter based on a $1.5m/s$ speed threshold and a $1g$ acceleration threshold are applied to the position estimation. The accuracy of the position estimation method was evaluated by comparing the estimated position of a person with its known ground path during 10 trials. Results show that this method is able to estimate the position of a person with an average error of $4.2cm$ (s.d. $2.6cm$) and a maximum error of $10.4cm$.

### C. Person Following

With the knowledge of the position of the person the robot can perform person following. Three algorithms were developed and implemented on our platform:

1) direction-following: the robot always goes in the direction of the person it follows, such as in [1].

2) path-following: the robot reproduces the path of the person. With this technique the robot can follow a person in an environment with low-height obstacles, e.g. chairs or tables, assuming that the path taken by the person is free of obstacles (see Fig. 2).

3) adaptive following: with an a priori obtained map of the environment the robot continuously computes and takes the shortest path to reach the position of the followed person, such as in [2].

### D. Position-invariant Dynamic Gesture Recognition

Gesture recognition employs the multi-dimensional dynamic time warping algorithm, similar to [3] and [4], which is used to align and compare two time sequences. In this work, coordinates of human joints, provided by the Kinect sensor, are used. In order for the algorithm to be independent from the relative speed, position and orientation of the robot and

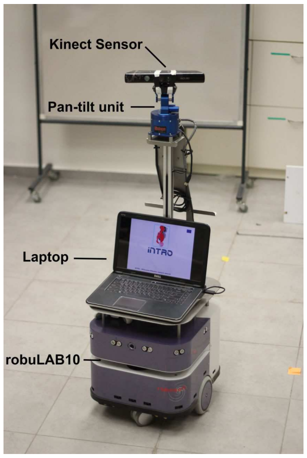


Fig. 1. The robotic platform.

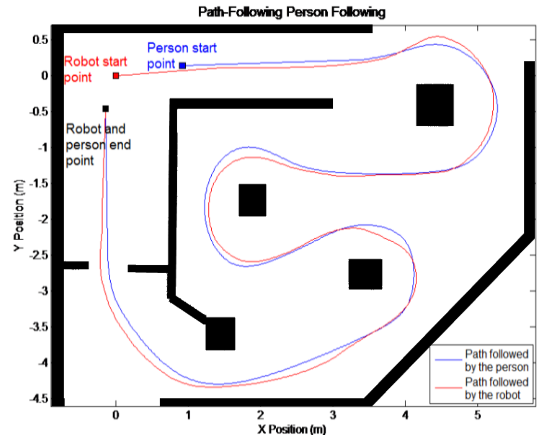


Fig. 2. Example of a path-following algorithm.

the person, the joints coordinates are transformed from the Kinect's frame of reference to the person's frame of reference. This frame of reference is defined by the center point of the left and right hip joints and a unit vector orthogonal to the ground. For a formal description of the algorithm and an evaluation of performance, please refer to [5].

### *E. Hardware*

The robotic platform is a Robosoft's robuLAB10, customized with a rigid structure, which supports a TRACLabs Biclops PTM (see Fig. 1). The mechanism can support a maximum payload of $4kg$, which was enough to carry the Kinect sensor. The robot, the Kinect sensor and the PTM were controlled by a laptop powered by an Intel quad-core i7 Q740 CPU with 4 GB of RAM.

## III. Case Study: Person-Following by a Robot Controlled by Gestures

To demonstrate capabilities of the described robotic platform, a person-following application was developed and tested, which was controlled by gestures. It used the specific capabilities of the platform: person tracking to keep the sensor oriented toward the person, person position estimation and path following to make the robot follow the person and dynamic gesture recognition to control the robot.

### *A. System Behavior*

Three robot states were defined: "passive tracking", "active tracking" and "following". To switch between the states four command gestures were created: "start active tracking", "stop active tracking", "start following" and "stop following". In "passive tracking", the PTM and the robot were static. In "active tracking", the robot remained static but the PTM was moving to direct the sensor toward the person. In "following", the robot followed the person using the path-following algorithm.

### *B. Experiments*

To test the systems performance, 10 trials were conducted. The user was instructed to signal to the robot to start to actively track them, then to follow them in an environment with obstacles and along a path similar to the one displayed on Fig. 2 and finally to stop the following and return to the "passive tracking" state. Performance was measured using four metrics: 1) task completion, 2) number of correctly recognized commands (hit), 3) number of incorrectly recognized commands (false positive) and, 4) number of undetected commands.

### *C. Results*

In all the trials, the user was able to initiate and to stop tracking and following using gestures with a goal of guiding the robot to the desired location. In 10 trials, out of 40 commands (4 commands per trial), only 5 were missed and had to be repeated and there were no false positives. These results demonstrate the feasibility of controlling the robot using gestures. In particular, it has been proven that it is possible to simultaneously track a person in motion, estimate their position and recognize gestures using a single depth sensor mounted on a mobile platform without the use of external sensors.


## Acknowledgment

This work has been financed by the FP7-ITN in the Marie-Curie People Programme: INTRO (INTeractive RObotics research network), grant agreement no.: 238486 and by the German Research Foundation within the GRK 1589/1.


## References


[1] K. Morioka, J.-H. Lee, and H. Hashimoto, "Human-following mobile robot in a distributed intelligent sensor network," *IEEE Trans. Ind. Electron.*, vol. 51, no. 1, pp. 229 – 237, feb. 2004.

[2] G. Doisy, A. Jevtić, E. Lucet, and Y. Edan, "Adaptive person-following algorithm based on depth images and mapping," in *Proc. of the IROS Workshop on Robot Motion Planning*, Oct. 2012.

[3] G. A. ten Holt, M. J. T. Reinders, and E. A. Hendriks, "Multi-dimensional dynamic time warping for gesture recognition," in *Proc. of the Conf. of the Advanced School for Computing and Imaging*, 2007.

[4] A. Corradini, "Dynamic time warping for off-line recognition of a small gesture vocabulary," in *Proc. IEEE ICCV Workshop on Recognition, Anal., and Tracking of Faces and Gestures in Real-Time Syst.*, ser. RATFG-RTS '01, 2001, pp. 82 –89.

[5] S. Bodiroža, G. Doisy, and V. V. Hafner, "Position-invariant, real-time gesture recognition based on dynamic time warping," in *Proc. of ACM/IEEE Int. Conf. on Human-Robot Interaction*, 2013, in press.