Human-Robot Collaboration using Micro-delegation and Gesture Recognition

Author Names Omitted for Anonymous Review. Paper-ID [add your ID here]

***Abstract*—We present a human-robot interaction system, capable of achieving high-level user requests through micro-delegation to humans. The robot – dubbed Niccolo – navigates in a human environment and completes tasks by seeking out humans and requesting assistance through minimal verbal interactions and gesture recognition. The system eventually achieves the high level request through a number of such interactions.**

I. Introduction

Can a robot find, buy, and bring you a bag of chips if it doesn't know what chips are or where to buy them, and it has no hands? Low-cost RGB-D sensors, for the first time, allow low-cost tracking and analysis of human skeletons. As such, human-robot interaction research, conventionally tied to expensive and high-tech laboratory set-ups, is now possible in battlefield conditions of real hallways, offices and workplaces. The goal is that of robotics researchers over four decades ago - to create a robotic platform that can accomplish high level tasks by collaborating with humans. We begin by hypothesizing that any human-realizable mission can be fulfilled by a robot if the robot exploits humans in a goal-oriented way. For example Bauer et al. [1] recently demonstrated a robot was able to find its way around a city by asking passers-by for directions. However, the system was implemented in an ad-hoc way with several pre-programmed routines of detecting people from 2D images. We propose to extend this concept by allowing the robot to pursue generic human-assigned missions utilizing the novel RGB-D sensor (such as the Microsoft Kinect) Through a process of micro-delegation, a complex high level task can be accomplished by relying on the help of humans in the immediate surroundings of the robot. The key low-level sensory human detection and tracking is accomplished through custom routines of learning skeleton gestures, such as pointing and walking.The robot receives a high-level structured English command of the type “bring me coffee”. The robot then attempts to achieve the goal by locating a nearby human and asking a clarification question, such as “where can I get coffee?” The robot identifies the direction by recognizing the humans gesture (pointing in the prescribed direction) and proceeding as suggested. Through a series of subsequent interactions the robot interacts with humans to achieve the high level goal.

II. Method

The robot — a Videre Erratic wheeled platform — sets out upon the world in search of a person. Once found, the robot asks the person if it can help them in any way. The task is input via the simple UI in the attached monitor. Figure 2 shows the first few interactions. Most communication to the person is done via synthesized speech, but the words spoken are shown on the screen as well for clarity.

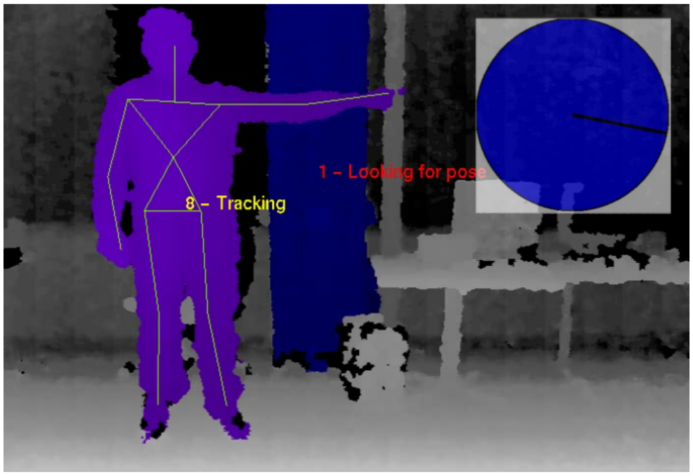
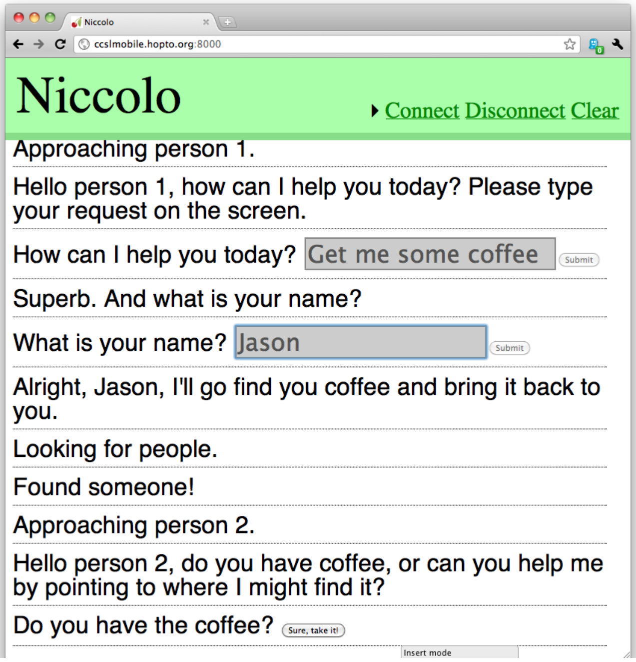


Fig. 1. Pointing gesture recognition and mapping.

After obtaining a task (currently limited to fetching objects) from the first user, the robot looks around for other people and, when detected, approaches them for assistance. The robot asks the people if they either have the goal object, or if not, if they could point in a direction where the robot would be more likely to find it. Figure 2 shows the interaction until the moment when the robot is waiting for the response from the second person, either by pressing the “Sure, take it!” button, indicating that they did have the desired object and have placed it on the robot's platform, or by pointing in the direction of the goal object.

A machine learning algorithm (SVM) is used atop the OpenNI skeletal tracking [2] to classify between gestures, such as passive standing and pointing. Skeletal tracking uses a single body calibration to track all humans, which has generalized fairly well and allows us to bypass the awkward “calibration pose” for every human encountered. The robot continues to explore the environment in this manner until the object has been obtained, at which point it switches states and finds its way back to the original user to deliver the goal object using a similar algorithm. Snapshots of several interactions are shown in Figure 4.



Niccolo

Connect Disconnect Clear

Approaching person 1.

Hello person 1, how can I help you today? Please type your request on the screen.

How can I help you today? Get me some coffee

Superb. And what is your name?

What is your name? Jason

Alright, Jason, I'll go find you coffee and bring it back to you.

Looking for people.

Found someone!

Approaching person 2.

Hello person 2, do you have coffee, or can you help me by pointing to where I might find it?

Do you have the coffee?

Fig. 2. Simple GUI shown on the robot's screen through which it augments verbal (output) and gesture (input) interaction with people it encounters.

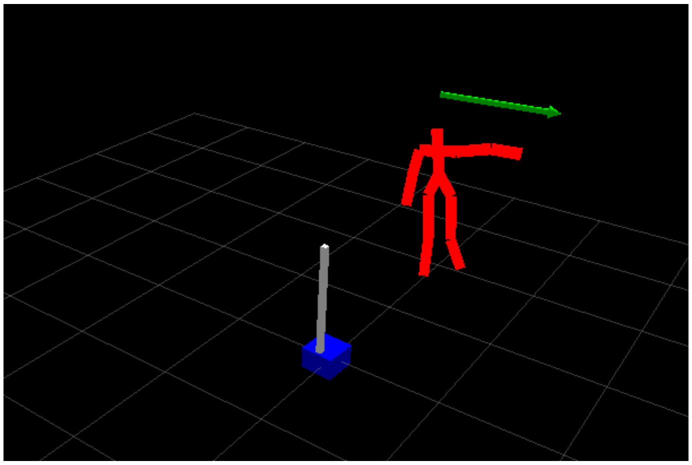


Fig. 3. RVIZ visualization of a user interaction through pointing.

III. RESULTS

An autonomous robot with a user interface allowing for basic interaction with a human (text-to-speech and speech-to-text) was deployed in the hallways of the building (Duffield Hall) housing the robotics laboratory of Cornell University. Through a sequence of interactions, the robot was successful at detecting humans and approaching them for assistance with a task (getting coffee). Detection and tracking of humans with over 85 percent accuracy in a real-world environment was instrumental towards realizing the human-assisted task completion.

Fig. 4. Several interactions. From top left: approaching first user, first user types in request for coffee, second user points toward where coffee is available, third user has coffee and offers it to robot.

REFERENCES

[1] Andrea Bauer, Klaas Klasing, Georgios Lidoris, Quirin Mhlbauer, Florian Rohrmller, Stefan Sosnowski, Tingting Xu, Kolja Khnlenz, Dirk Wollherr, and Martin Buss. The autonomous city explorer: Towards natural human-robot interaction in urban environments. *International Journal of Social Robotics*, 1:127–140, 2009. ISSN 1875-4791. URL http://www.lsr.ei.tum.de/fileadmin/publications/ace_soro.pdf. 10.1007/s12369-009-0011-9.

[2] Open NI Team. Open ni website. http://www.openni.org/, 2011.