Integrating Multimodal Interaction and Kinesthetic Teaching for Flexible Human-Robot Collaboration

R. Caccavale[1], A. Finzi[1], D. Lee[2], E. Leone[1], S. Rossi[1], M. Saveriano[2] and M. Staffa[1]

Abstract— **We present a human-robot interaction framework that integrates multimodal collaborative execution and kinesthetic teaching to accomplish complex and new collaborative tasks in an industrial scenario. We consider the case of a hospital scenario where a human user interacts with a robotic arm in order to collect and to arrange tools.**

I. Introduction and System Overview

An effective cooperation between a human being and a robotic co-worker during the execution of complex tasks requires natural interaction and continuous and incremental adaptation. In this paper, we present an approach to human-robot collaboration that combines multimodal interaction and kinesthetic teaching. The aim is to allow a human operator to continuously and naturally switch from collaborative execution to teaching modality and vice versa by exploiting multimodal inputs -mainly speech and gestures- and physical interaction. The multimodal interaction system recognizes multiple human commands/actions [1] providing an interpretation of users intentions according to the context. The interpretation process is based on a late fusion approach: the results of classifiers of the single modalities - gesture (LDCRF-based recognition) and speech (Julius recognizer) - are integrated by the fusion engine (exploiting probabilistic context-free grammars), while a dialogue manager accomplishes the semantic interpretation of the observations according to the interaction context (see [2], [3]). Kinesthetic teaching requires a low level control that guarantees a safe physical interaction and an easy guidance [4]. In this work, we use the gravity compensation control to have an ideally massless robot that the user can easily and safely guide.

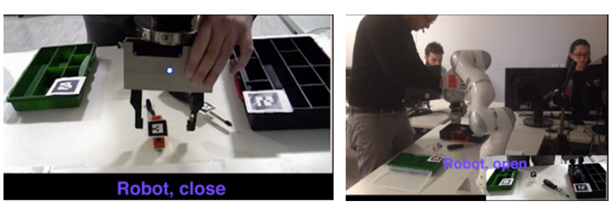

Fig. 1. Hospital Scenario: kinesthetic teaching and speech-based interaction.

*The research leading to these results has been supported by the SAPHARI project (FP7/2007-2013 under grant agreement ICT-287513).

[1]DIETI, Università degli Studi di Napoli "Federico II" via Claudio 21, 80125, Naples, Italy.

[2]LSR, Technische Universität München, Karlstrasse 45, 80333, Munich, Germany.

II. Case Study: DLR Hospital Scenario

The hospital domain concerns the process of quality control, preparation, and packaging in a hospital center. In this context, the operator and robot co-worker are involved in the task of sorting instruments on a tray. The task is composed of 5 steps: the human brings a tray of unsorted instruments to the robot; he/she checks the instruments and puts them within the robot workspace; the human asks the robot to help; during the interaction the robot manipulates specific objects while the human checks the instruments and orders the other objects; the task is completed when the tray contains ordered and checked instruments.

a) Set-up: As a set-up, we considered a KUKA LWR IV+ manipulator endowed with a gripper operating in a workspace monitored by two kinect cameras, one for user interaction and another used to track the objects (recognized via the qr code) on the table. The setting includes two trays and three objects: a screwdriver, a yardstick (the unknown object), a tape (see Fig. 1).

b) Execution and Teaching modes: We assume the robot already aware about the target location of some tools, while the human should instruct the robot where to place (which box) the other objects thought kinesthetic teaching. During the interactive execution the user can communicate a set of primitive intentions (*Point*, *Take/Leave*, *Give*, *No*, *Find*, *Come*, *Stop*, *Switch*). The switch can be always invoked to start a teaching session, in this case the robot goes in gravity compensation waiting for the human physical guidance. In this mode, the user intentions are suitably interpreted according with the novel interaction schema. The operator first shows the target object to the robot (kinect), then he/she moves the robot arm from the object position to the target (kinesthetic teaching). In this phase additional commands can be provided (e.g. *Open/Close* gripper), mainly vocally, since the human physically interacts with the robot. Before stopping the teach mode, the human can ask the robot to repeat the learned sequence to verify its execution. Tests with the proposed framework are currently in process (see [5]).

References

[1] S. Iengo, S. Rossi, M. Staffa and A. Finzi. "Continuous gesture recognition for flexible human-robot interaction," ICRA, 2014.
[2] S. Rossi, E. Leone, M. Fiore, A. Finzi and F. Cutugno. "An Extensible Architecture for Robust Multimodal Human-Robot Communication," IROS, 2013.
[3] L. Lucignano, F. Cutugno, S. Rossi and A. Finzi. "A Dialogue System for Multimodal Human-Robot Interaction," ICMI, 2013.
[4] M. Saveriano, S. An and D. Lee. "Incremental Kinesthetic Teaching of End-Effector and Null-Space Motion Primitives," ICRA, 2015.
[5] Video available: https://youtu.be/nEOWfVMZHzc