Supervised Hand-Guidance during Human Robot Collaborative Task Execution: a Case Study

Jonathan Cacace[a], Riccardo Caccavale[a] and Alberto Finzi[a]

[a]DIETI, Università di Napoli Federico II, via Claudio 21, 80125 Napoli, Italy

Abstract

We present and discuss a human-robot collaboration system suitable for supervising the execution of structured manipulation tasks in industrial assembly scenarios. As a case study, we consider the application domain proposed in the context of the project (PON R& I 2014-2020) ICOSAF (Integrated collaborative systems for Smart Factory) in which a human operator physically interacts with a collaborative robot (Cobot) to perform multiple item insertion tasks in a shared workspace. The proposed system combines hierarchical task orchestration and human intention recognition during human-robot interaction through hand-guidance. We provide an overview of the system discussing an initial experimental evaluation.

Keywords

Human-Robot Collaboration, Hand-guidance, Intention recognition, Flexible Manufacturing

1. Introduction

Collaborative robotic systems (CoBots) enable humans and robots to safely work in close proximity during the execution of shared tasks [1] merging their complementary abilities [2]. In this work, we address these issues considering an industrial assembly scenario in which a human operator interacts with a lightweight robotic manipulator through hand-guidance in order to accomplish multiple and structured operations in a shared workspace. Specifically, in the proposed application domain a CoBot should support a human worker during the insertion of accessories into carbon-fiber monocoque cells for car production. This case study is provided by the Italian project (PON R& I 2014-2020) ICOSAF (Integrated collaborative systems for Smart Factory), whose aim is the design and development of models and methods for collaborative factories. In order to accomplish these tasks, we propose to deploy a collaborative human-robot interaction system that combines task supervision and orchestration with continuous interpretation of the human physical guidance. In the proposed framework, the intentions conveyed with the operator physical interventions are interpreted with respect to the planned activities and motions, while the robot behavior is suitably adapted by switching tasks, changing targets, adjusting trajectories, and regulating the robot compliance to the human guidance. Flexible task execution and fluent human-robot interaction are supported exploiting the ex-

---

The 7th Italian Workshop on Artificial Intelligence and Robotics (AIRO 2020), November 26, Online.

jonathan.cacace@unina.it (J. Cacace); riccardo.caccavale@unina.it (R. Caccavale); alberto.finzi@unina.it (A. Finzi)

0000-0002-1639-5655 (J. Cacace); 0000-0001-8636-7628 (R. Caccavale); 0000-0001-6255-6221 (A. Finzi)

ecutive framework proposed in [3, 4, 5], which exploits supervisory attention and contention scheduling [6, 7] to monitor human behaviors and suitably orchestrate multiple hierarchically structured tasks. In particular, supervisory attention permits to smoothly integrate plan guidance and human guidance through top-down and bottom-up regulations. Collaborative task execution is also affected by the interpretation of the human guidance with respect to the supervised activities. Following the approach by [8, 9], depending on the operational state, the supervisory system enables possible subtasks, targets and trajectories, which are continuously evaluated by intention recognition processes. Each possible trajectory is assessed by a LSTM network that infers the intention of the operator to follow/contrast the manipulator motion towards a target point, deviate from the latter, or use the robot manipulator in direct manual control. In this paper, we provide an overview of the system at work in the industrial assembly case studies describing an experimental setup used to perform an initial assessment of the proposed framework.

2. Case Study

We consider an industrial assembly case study, proposed in the context of the Italian project (PON R& I 2014-2020) ICOSAF (Integrated collaborative systems for Smart Factory) project, whose aim is the design and development of models and methods for collaborative factories. In the assembly operative scenario, a CoBot should support human workers during the insertion of accessories into carbon-fiber monocoque cells for car production. Specifically, we focused on the task of inserting metallic items (called bigheads) on the monocoque, which is usually manually executed. Human-robot collaboration is particularly suitable for this task since both the human dexterity and the robot precision are required. Indeed, the main requirements are: positioning accuracy, repeatability, supervision systems for the application of the correct item, safety, improvement of the operator's ergonomics. During task execution, both the human and the robot should be able to work on the monocoque, hence concurrent insertion of items should be possible. The operator should always be allowed to hand guide the robot to point the end-effector towards a different target or a different working area.

3. System Architecture

The architecture of the human-robot collaborative system is illustrated in Fig. 1. The *High-Level Control System* (HL) is responsible for task generation, supervision and execution, while the *Low-Level Control System* (LL) manages the execution of the trajectories proposed by the high-level system integrating the human physical guidance. Task supervision and orchestration relies on the attention-based executive framework proposed in [3, 4, 5]. During task execution, the human operator can physically interact with the CoBot (force/position feedback) and these interventions are simultaneously interpreted at the different layers of the architecture. Depending on the task, the environmental context, and the human interventions, the Executive System (top-down) retrieves hierarchical tasks in the Long Term Memory (LTM) and allocates them in the Working Memory (WM) (see Fig. 2). The primitive operations are associate with behaviors/processes that compete for the execution (*Behavior-based System*). In

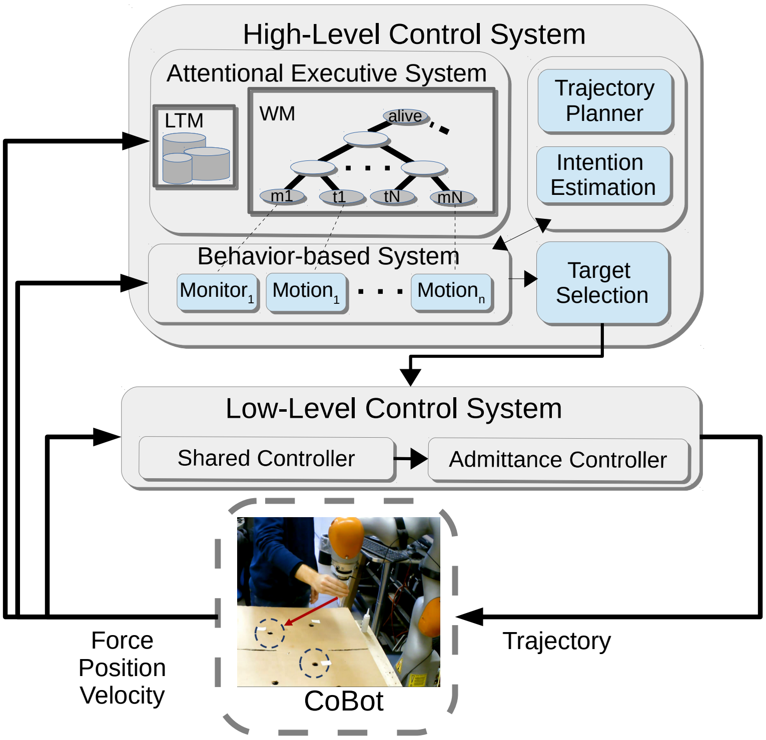

Figure 1: System Architecture. The *High-Level Control System* manages task generation, supervision and execution; the *Low-Level Control System* (LL) manages compliant trajectory execution.

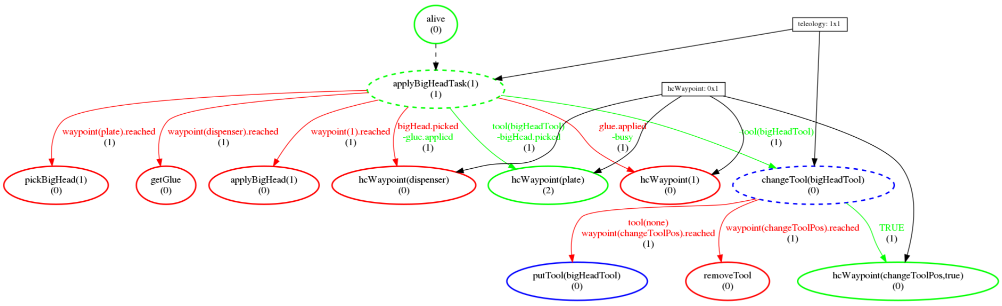

Figure 2: Tasks allocated in working memory during human-robot collaboration. Red, green, and blue ovals are for disabled, enabled, and executed activities, respectively.

this case, motion behaviors (e.g., *pick*, *place*) are associated with target positions and a trajectories (generated by the *Trajectory Planner*), while each possible trajectory is continuously monitored in order to estimate (*Intention Estimation*) the one more aligned with respect to the human guidance. Intention estimation relies on LSTM networks, one for each trajectory, that classify the operator interventions as *aligned*, *deviating*, *opposed*, *opposed deviating*. The classification results (along with the top-down attentional regulations provided by the WM) are then exploited to influence the CoBot selection of targets and trajectories (*Targret Selection*).

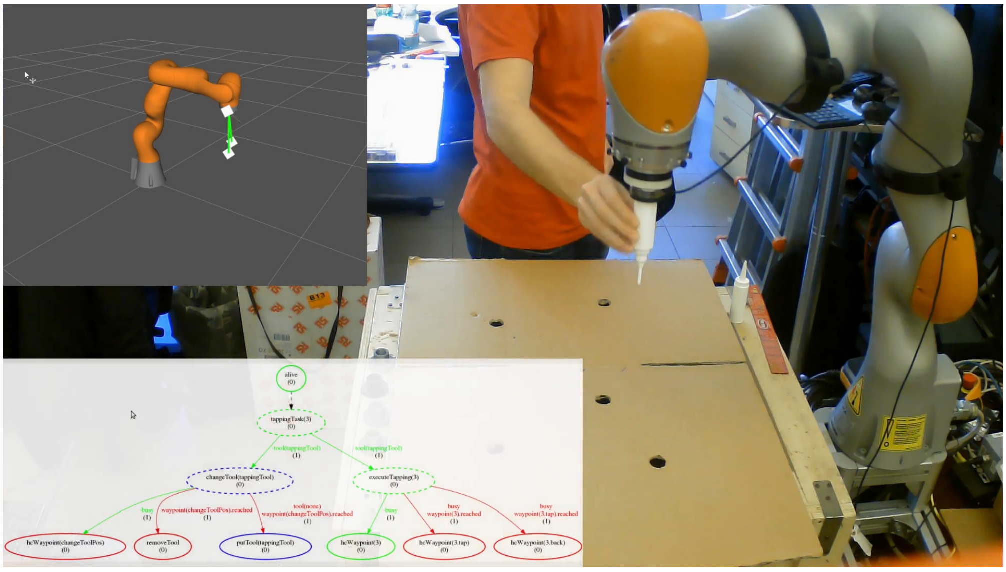

Figure 3: Experimental setup. The operator can hand-guide the end-effector towards the desired target points while the supervisory system interprets the human interventions in the context of the task allocated and enabled in WM (bottom-left rectangle).

Finally, the lowest level of the architecture implements a *Shared Controller* aimed at mixing the inputs generated by the the human operator (*Shared force*) and the ones needed to perform robot motion. An (*Admittance Controller*) integrates the human and the robot guidance.

4. Experimental setup

In order to test the proposed human-robot collaboration system, we implemented the test-bed depicted in Fig. 3, which illustrates a mockup representing a surface with positions where operations must be executed. Since our main interest was to test human-robot collaboration, we considered only end-effector movements towards target positions, while the final insertion operations were simulated. We deployed the Kuka LBR IIWA manipulator, controlled via ROS middleware running on a standard version of Ubuntu 18.04 GNU/Linux OS. The ATI Mini 45 Force sensor has been used to detect the human input to command the robot. As for the LSTM network, it has been implemented using TensorFlow[1] library through Keras[2].

In this setting, we focused on the system ability in following the operator hand-guided operations considering both quantitative (effort and execution time) and qualitative (NASA-TLX [10] inspired questionnaire) assessments. The task assigned consisted in the execution of multiple simulated operations (5 ordered item insertions) on target positions (see Fig. 3). Since the operations' sequence is provided to the operator, but this is unknown to the robot, the human co-worker is requested to continuously correct the manipulator trajectories via hand-guidance

---

[1] https://www.tensorflow.org

[2] https://keras.io

to obtain the requested order of execution. We compared three interaction modalities: proactive, guided, and passive. In the proactive mode, the robot directly heads to the next target without human confirmation, while in the guided mode the robot always waits for an operator physical input to reach the next target. The passive mode is used as a baseline; in this case the robot is fully compliant with respect to the human physical guidance. The experiment was carried out by 40 testers (graduate or post-graduate students). Preliminary results in this experimental setting show the advantage of the proactive and guided modalities with respect to the passive one, with a slight preference for the guided mode. Specifically, as expected, we observed that both guided and proactive modes significantly reduce the operator physical effort (measured as the cumulative impulse applied to the robot) with respect to the passive mode, with less effort measured in the proactive mode (we observed an average effort of 29, 52, 97 Ns for the proactive, guided and passive modes, respectively). On the other hand, qualitative evaluation results show that guided mode was considered as more reliable, readable and satisfactory; users also estimated less pressure and mental demand. In this respect, the guided mode seems preferred to the proactive mode because it provides the users with more control during task execution.

5. Conclusion

We presented a human-robot collaboration system for CoBots that supports hand-guided human interventions during the execution of structured tasks in an industrial assembly scenario. The proposed system combines hierarchical task orchestration and human intention recognition during physical interaction. We described the application domain and the system design discussing different interaction modes (passive, guided, proactive). The collected results show the advantage of the proposed assisted modalities with respect to the passive one. We are currently designing and developing tasks for the real industrial workspace considering issues like positioning accuracy, task reliability, safety, and ergonomics. We are also investigating adaptive multimodal extensions of the framework [11, 12] including additional communication channels [13] (e.g. gestures and speech) along with associated fusion methods [14] and adaptive interfaces [15].

Acknowledgments

The research leading to these results has been partially supported by the projects REFILLs (H2020-ICT-731590) and ICOSAF (PON R& I 2014-2020).

References

[1] J. A. Corrales, G. J. Garcia Gomez, F. Torres, V. Perdereau, Cooperative tasks between humans and robots in industrial environments, International Journal of Advanced Robotic Systems 9 (2012). doi:10.5772/50988.

[2] D. Romero, P. Bernus, O. Noran, J. Stahre, Å. Fast-Berglund, The operator 4.0: human cyber-physical systems & adaptive automation towards human-automation symbiosis

work systems, in: IFIP international conference on advances in production management systems, Springer, 2016, pp. 677–686.

[3] R. Caccavale, A. Finzi, Flexible task execution and attentional regulations in human-robot interaction, IEEE Transactions on Cognitive and Developmental Systems 9 (2017) 68–79.

[4] R. Caccavale, M. Saveriano, A. Finzi, D. Lee, Kinesthetic teaching and attentional supervision of structured tasks in human–robot interaction, Autonomous Robots 43 (2019) 1291–1307. doi:10.1007/s10514-018-9706-9.

[5] R. Caccavale, A. Finzi, Learning attentional regulations for structured tasks execution in robotic cognitive control, Autonomous Robots 43 (2019) 2229–2243.

[6] D. A. Norman, T. Shallice, Attention to action, in: Consciousness and self-regulation, Springer, 1986, pp. 1–18.

[7] R. P. Cooper, T. Shallice, Hierarchical schemas and goals in the control of sequential behavior, Psychological Review 113 (2006) 887–916.

[8] J. Cacace, R. Caccavale, A. Finzi, V. Lippiello, Interactive plan execution during human-robot cooperative manipulation, IFAC-PapersOnLine 51 (2018) 500 – 505. URL: http://www.sciencedirect.com/science/article/pii/S2405896318332919. doi:https://doi.org/10.1016/j.ifacol.2018.11.584, 12th IFAC Symposium on Robot Control SYROCO 2018.

[9] J. Cacace, A. Finzi, V. Lippiello, Enhancing shared control via contact force classification in human-robot cooperative task execution, in: F. Ficuciello, F. Ruggiero, A. Finzi (Eds.), Human Friendly Robotics, Springer International Publishing, Cham, 2019, pp. 167–179.

[10] S. Hart, L. Staveland, Development of nasa-tlx (task load index) : Results of empirical and theoretical research, Human Mental Workload. Advances in Psychology (1988) 139–183.

[11] R. Caccavale, E. Leone, L. Lucignano, S. Rossi, M. Staffa, A. Finzi, Attentional regulations in a situated human-robot dialogue, in: The 23rd IEEE International Symposium on Robot and Human Interactive Communication, IEEE RO-MAN 2014, Edinburgh, UK, August 25-29, 2014, 2014, pp. 844–849.

[12] R. Caccavale, J. Cacace, M. Fiore, R. Alami, A. Finzi, Attentional supervision of human-robot collaborative plans, in: 2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), 2016, pp. 867–873. doi:10.1109/ROMAN.2016.7745221.

[13] S. Iengo, S. Rossi, M. Staffa, A. Finzi, Continuous gesture recognition for flexible human-robot interaction, in: 2014 IEEE International Conference on Robotics and Automation, ICRA 2014, Hong Kong, China, May 31 - June 7, 2014, 2014, pp. 4863–4868.

[14] J. Cacace, A. Finzi, V. Lippiello, A robust multimodal fusion framework for command interpretation in human-robot cooperation, in: 26th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN 2017, Lisbon, Portugal, August 28 - Sept. 1, 2017, 2017, pp. 372–377.

[15] J. Cacace, R. Caccavale, A. Finzi, V. Lippiello, Attentional multimodal interface for multidrone search in the alps, in: 2016 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2016, Budapest, Hungary, October 9-12, 2016, 2016, pp. 1178–1183.