Review of the Doctoral Thesis „Human-Robot Interaction: Advanced Task-Centered Interfaces for Non-Expert Users“ Written by Zdeněk Materna

Reviewer: doc. Ing. Zdeněk Míkovec, Ph.D., Czech Technical University in Prague

The presented thesis is dealing with the problem of designing human-robot interaction (HRI) for controlling and programming of collaborative robots by non-expert users. This topic is very urgent as robots become a part of our everyday life. With the increasing complexity of robots and tasks robots are solving, increasing number of situations where the close and collaborative HRI takes place, we are facing problems that were not solved properly yet. The state of the art analysis (chapter 3 and corresponding subsections in chapters 5-8) shows a significant number of attempts and partial solutions focused on the problem of controlling and programming a robot. However, these solutions seem to be unsatisfactory for non-expert users. For a better understanding of the limitations and challenges in this particular field of HRI, it would be beneficial to present some classification in the introduction part of the thesis and conclusions crafting the problem space that will be tackled in the thesis at the end of the state of the art chapter 3.

In the thesis, we can recognize the following scientific contributions with potentially high research impact:

1. Introduction of task-centered interaction concept, which integrates following interaction approaches: interaction within the task space, interaction on a higher level of abstraction, and context-sensitive interaction.
2. Detail report on two implemented and tested systems, which user interfaces were designed according to the presented task-centered interaction concept.

The research results and scientific contributions are described in chapters 5, 6, 7, and 8.

In chapter 2, there are formulated six research objectives that demonstrate the focus of the thesis. However, there are some statements that require deeper explanation.

In objective no. 3, there is a hypothesis formulated, that by the introduction of triggering and parametrization of robot’s autonomous functions would help to keep the mental workload low and thus maintain the efficiency of collaboration.

> *To understand this objective more clearly, is it possible to classify the autonomous functions, describe what are the main contributors to the mental workload, and what are the parameters of collaboration efficiency?*

In objective no. 5, there is stated, that the capacity of short-term memory (STM) is limiting the human performance in the context of a joint task.

> *Is there any example or observation of the negative effect of STM on the performance of a joint task, that can be reported?*

In objective no. 6, there is a rather vague definition of a successful evaluation of the designed user interfaces "..., if the main aim of the interface will be satisfied without major issues, it could be claimed that the objective was fulfilled."

> *Is it possible to define more concrete evaluation metrics, where the terms like "satisfied", "major issues", "fulfillment of objective" will be defined?*

Chapter 4 describes in detail the proposed concept of task-centered interaction (TCI), which seems to be complex with the potential to solve the issues mentioned earlier in this thesis. Five key characteristics can be identified here:

TCI1: interaction elements embedded into the scene
TCI2: use of robot capabilities
TCI3: communication of robot inner state
TCI4: context-sensitive UI "filtering"
TCI5: task-appropriate (interaction) modalities

Each characteristic itself is clearly defined, and concrete examples of their application are provided (chapter 4.2).

> *However, is it possible to define in a formal and more general way how these key characteristic of the TCI concept should be combined, based on the experience with the implementation of two presented systems in chapter 5, 6, and 8?*

In chapter 5, the user interface for teleoperating assistive robots is introduced. The final user interface design implements TCI1 and TCI2 characteristics of the TCI concept. The 3D voxel-based environment model seems (according to very extensive evaluation with 81 participants in 5 design steps) to be a very successful approach (better than a geometric mapping) to provide a user with an overall information about the environment, which helps with better orientation and teleoperation of the robot's base and the robotic arm. Showing live RGB-D data of the current field of view inside a 3D map is a very interesting approach with high potential to help users with orientation in case of a very small field of view or rich and dynamically changing environments.

Several functions are not described (e.g., leading the user through sub-tasks – giving instructions, an image of the whole UI including all components and controls) or described very briefly (direct driving, etc.). Some details can be found in [88 Fig. 7]. However, the main text of the thesis should be self-containing.

The usage of two pointing devices (2D and 3D mouse) is an interesting approach. More details on the motivation and issues of usage of such a setup would be of high benefit.

> *Why there are introduced two pointing devices? Would it not be sufficient to use only one of them and which one? Where there observed any issues related to the necessity of the user to switch between devices and map operations to the right one (time delay of hand movements between devices, learning and time issues with the mental mapping of operations to a particular device)? What about considering interaction with both devices in parallel (setting the view pint with a 2D mouse and teleoperating the robot's base with a 3D mouse)?*

In chapters 6 and 8, the user interface for industrial robot programming is introduced. There were two experiments conducted. First one (chapter 6) focused on interaction errors and multimodality of interaction. The UI implemented for this purpose covered TCI1, TCI2, and TCI5 characteristics of the TCI concept. The experiment was very complex. The experiment design was 3x4x5 (error x tasks x modality). There were three user groups divided according to the frequency of errors. The order of tasks was fixed, and modality was ordered randomly. Totally 39 participants took part in the experiment. As the best modalities gesture and 6D pointing device were identified, which were also the fastest. Direct robotic arm programming failed in both measured parameters. These results bring very useful insight into the relationship between errors, tasks and interaction modalities.

Is it possible to define some general guidelines for designing user interfaces for similar robots in collaborative environments, that can be applied to slightly different setups (variations of the tasks, robot capabilities, modalities at disposal, environment features)?

According to the experiment design, the statistics on completion time should be reported differently, using repeated measures ANOVA and Post hoc (e.g., Scheffe) test for determining significant differences between each pair of modalities.

The second experiment focused on the user experience evaluation of collaborative robot programming (chapter 8). According to a rather smaller number of participants (6), this was more qualitative evaluation. The user interface designed covered TCI1, TCI2, TCI3, and TCI5 characteristics of TCI concept. The SUS evaluation ranges from 62.5 (C-), which is rather under the average and 87.5 (A), which is far above the average. This seems to be a promising result demonstrating clearly the potential of the TCI concept. However, according to the high spread of SUS values, further investigation is highly recommended.

Is it possible to explain the high spread of SUS values? What caused the low evaluations and what caused the high ones?

As the number of participants is small reporting of means and standard deviations is not recommended. Moreover, the SUS results should be reported using percentile ranking [zm1]. It seems that the TLX questionnaire is not properly used here (according to its complexity, it should be used after each task rather than after the test [zm2]).

It would be interesting to see a comparison of various interaction setups to deeper investigate the differences.

How the effect on mental (cognitive) load, and attention switching was measured (stated in chapter 8.7)?

There were moderator interventions reported in the experiment. How could these interventions influence the success rate and the completion times of the experiment sessions?

The TCI concept was demonstrated except the TCI4 characteristics.

Is it possible to demonstrate the implementation of TCI4 characteristic (context-sensitive UI "filtering") on some user interface implemented within the frame of the thesis?

The presented results are of high scientific level and were correspondingly presented on relevant scientific conferences and published in international journals. There appeared already citations of the key scientific papers. The h-index of the candidate is according to citation databases 4/4/6 (WoS/Scopus/Google Scholar).

Conclusion: The thesis and present publication results demonstrate that the candidate achieved valuable scientific results. According to The Higher Education Act (Law Nr. 111/1998), I do recommend the thesis for the presentation and defense with the aim of receiving a Ph.D. degree.

Prague 11. 9. 2019

doc. Ing. Zdeněk Míkovec, Ph.D.

Faculty of Electrical Engineering
Czech Technical University in Prague

[zm1] Sauro J. Measuring usability with System usability scale, 2011, https://measuringu.com/sus/

[zm2] Laubheimer P. Beyond the NPS: Measuring Perceived Usability with the SUS, NASA-TLX, and the Single Ease Question After Tasks and Usability Tests, 2018, Nielsen Norman Group, https://www.nngroup.com/articles/measuring-perceived-usability/