Online Intention Recognition in Computer-Assisted Teleoperation Systems

Nikolay Stefanov, Angelika Peer, and Martin Buss

Institute of Automatic Control Engineering, Technische Universität München
Theresienstrasse 90, 80333 München, Germany
{nikolay.stefanov, angelika.peer, mb}@tum.de
http://www.lsr.ei.tum.de

Abstract. Limitations of state-of-the-art teleoperation systems can be compensated by using shared-control teleoperation architectures that provide haptic assistance to the human operator. This paper presents a new approach for computer-assisted teleoperation, which recognizes human intentions and dependent on the classified task activates different types of assistances. For this purpose, time series haptic data is recorded during interaction, passed through an event-based feature extraction, and finally used for task classification by applying a Hidden Markov Model approach. The effect of the assistance function on human behavior is discussed and taken into account by training multiple classifiers for each type of assistance. The introduced approach is finally validated in a real hardware experiment. Results show an accurate intention recognition for assisted and non-assisted teleoperation.

Keywords: teleoperation, human intention recognition, shared control, computer assistance

1 Introduction

Despite recent advances in the design and control of haptic teleoperation systems, they are still characterized by a high execution time, failure rate, and low performance when compared to direct human manipulation. Because of hardware or software limitations existing teleoperation systems cannot satisfy objectives like robustness, transparency, performance, and high degree of presence at the same time. Providing assistance when executing a teleoperation task is considered a valuable strategy to overcome some of these limitations. To provide, however, a suitable assistance at the right moment in time, knowledge about the actual executed task is required. Thus, an automated, online intention recognition algorithm must be implemented, which is able to recognize and classify tasks at the moment in time they are executed by the human operator. While reliable classifications are achievable in case no assistance is provided, the provision of assistance leads to a drastic change of human behavior which again results in a high probability of misclassification. A reliable classification for both, computer-assisted as well as direct teleoperation is, however, a prerequisite for a correct activation of the assistance function. In this paper we present an intention recognition algorithm which shows a high recognition rate for both types of teleoperation.

1.1 Related Work

Haptic assistants are well known in the field of autonomous robot assistants where the scenario of collaboratively carrying an object has been studied intensively. Corteville et al. [1] e.g. proposed a robot assistant that actively supports the human while transporting an object from a given start to a prespecified target position. Their approach clearly goes beyond classical leader-follower architectures by implementing a robot that estimates human intentions from haptic signals and that contributes actively to the collaboration. Unterhinninghofen et al. [2] suggest a shared control assistance for bilateral teleoperation. The approach infers human intentions, predicts the trajectory of the operator, and based on this information applies corrections on the position signal at the remote side as well as activates forces of a virtual potential field at operator side.

Although the importance of identifying human intention and distinguishing between tasks is mentioned in both studies, Corteville et al. do not approach this problem online, but rather use methods for offline task segmentation and Unterhinninghofen et al. concentrates on a single task only such that no task classification is required.

Human intention recognition received more attention in fields like robot imitation, programming-by-demonstration, see [3], [4] and [5], or human-robot skill transfer [6] where mostly motion signals are studied and human intentions are represented on different levels of abstraction [7]. These kind of studies, however, do not consider any contact between human and robot and thus, intention recognition is unaffected by the actual robot behavior. This is not the case for haptic assisted teleoperation where human and assistant closely interact with each other.

In this work, we present an assistance control architecture which i) online [1] classifies the actual task by analyzing haptic data and ii) provides different types of haptic assistances. The presented approach takes into account changing human behavior in case assistance is applied and reaches a good task classification for both, computer-assisted as well as direct teleoperation.

2 Control Architecture for Computer-Assisted Teleoperation

This section provides an overview of the implemented control architecture for computer-assisted teleoperation. As illustrated in Fig. 1 a) a classical bilateral teleoperation architecture (solid line) is extended by introducing new building

[1] An **online** algorithm receives its input incrementally and responses to each input portion.

blocks for the computer assistance (dashed line). Since human operator and assistant share the control over the telerobot, a classical shared control architecture is realized.

2.1 Main Building Blocks of Computer Assistant

A detailed illustration of the computer assistant is given in Fig. 1 b). Its building blocks are briefly discussed in the following paragraphs. It generally consists of two main components, intention recognition and assistance function. The intention recognition block again contains two main building blocks, feature extraction and classifier. For a detailed description of these last two components please refer to our previous work [8]. Here only a short summary will be provided.

Feature extraction: The feature extraction algorithm uses haptic data as input signals and generates a highly compressed one-dimensional, discrete observation vector. The algorithm is event-triggered and emits a symbol if the input data crosses a certain threshold. This has a big advantage over sampling with fixed sampling rate as the resulting observation sequence is time-warping invariant and robust against noise and small variations in the human behavior ([8]).

Classifier: Since human intentions cannot be measured directly, but can be estimated by observing human behavior by means of a sequence of output symbols [9], multiple discrete Hidden Markov Models (HMM) are used for online intention recognition. For each HMM (each represents a certain task) the forward probability is computed and thus the first canonical problem over HMMs is solved: Given the HMM λ and an observation sequence $O = \{o_1, o_2, \ldots, o_\tau\}$ with $o_i \in V$, the probability that O is produced by λ is calculated. Finally, the HMM with the highest probability is assumed to have emitted the observed sequence and to represent the task the human is going to perform, respectively her/his intention.

Assistance function: Assistance functions are highly task dependent and often introduced to enhance task performance in terms of precision or task completion time. Assistance can either concentrate on a single modality by providing visual, auditory, or haptic feedback or be multimodal when providing a combination of these. As the focus of this paper lies on the haptic modality, only haptic assistance functions suitable for the usage in shared control architectures are considered.

2.2 Intention Recognition in Computer-Assisted Teleoperation Systems

When activating the assistance function the human behavior changes, which leads to new patterns in haptic signals measured during execution of the task.

Accordingly, the observation sequences generated by the feature extraction module change as well and thus, the previously trained classifier is not longer able to correctly classify these sequences. On this account, the influence of the assistance function has to be taken into account in the design of the intention recognition algorithm. A natural approach would be to compensate the influence of the as-

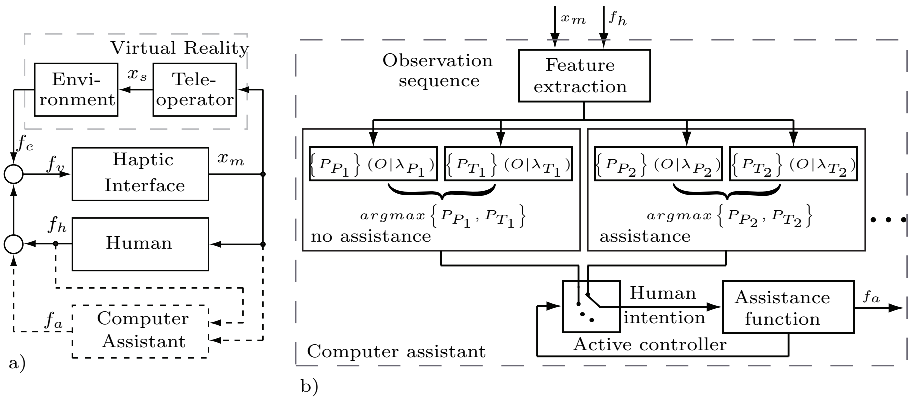

Fig. 1. Schematic view of: a) Control architecture for computer-assisted teleoperation; b) Computer assistant

sistance function on the recorded signals before passing them to the intention recognition algorithm. This could be achieved by applying the inverse of the assistance function on the recorded data. This approach assumes that the adjusted haptic signals used as input signals for the feature extraction block would have the same form as in the non-assisted case. Unfortunately, however, assistance functions affect the whole human interaction behavior and thus, patterns change significantly for the assisted and non-assisted case, compare e.g. Fig. 3. Consequently, when applying assistances the observed oberservation sequence contains symbols that are arranged in a different way than in the non-assisted case, which makes the classification fail. Finally, such an approach would also assume that the assistance function can be modelled mathematically, which can be difficult in certain cases. On this account, we follow an alternative approach to compensate the effect of the assistance function on the intention recognition. Instead of compensating the influence of the assistance function on the extracted features, we switch between different classifiers. For the non-assisted case as well as for each type of assistance function an own classifier has to be trained, see Fig. 1 b) for the case of only one assistance function. Since the information about the currently active assistance function is known, this information can be used to switch between the single trained classifiers. The drawback of this solution is obvious: it assumes that the task can still be performed even when having activated the wrong assistance function, because such training trials are necessary to train the single classifiers.

3 Experimental Evaluation

Scenario: For evaluation of the presented approach of online intention recognition, a scenario consisting of typical point-to-point movements is chosen, which requires repeated transferring of an object from a starting to a target position, see Fig. 2 a). In doing so, the subject applies a force f_h to overcome the inertia of the object and operates against friction f_r. The exercise consists of two randomly concatenated tasks: i) the transportation task (transportation of the object from the starting to the target position) and ii) the positioning task (positioning the object at the target location). Instead of using a teleoperation system with a real slave device, we used a haptic interface to interact with a virtual, haptically rendered object. In doing so, we are able to avoid artifacts originating from the communication channel and can easily perform experiments with various objects and environment dynamics.

To simulate the virtual object a haptic rendering algorithm was used which implements a rigid object with a mass of 3 kg that can be slided horizontally over a rigid ground. In order to increase the realness of the simulation, the following friction model was implemented:

$$f_f = f_{st} + f_k + f_v \quad \text{with} \quad f_{st} = \begin{cases} -f_h & \text{for } f_h < \mu_s f_N \\ 0 & \text{for } f_h \geq \mu_s f_N \end{cases}, \quad f_k = \mu_k sgn(f_h), \quad f_v = \mu_v \dot{x},$$

where f_{st} is the static, f_k the kinetic and f_v the viscous friction and $\mu_s = 0.5$, $\mu_k = 0.2$ and $\mu_v = 0$ are the corresponding friction coefficients. The human interaction force is denoted with f_h, the normal force with f_N. No visual feedback of the object was provided.

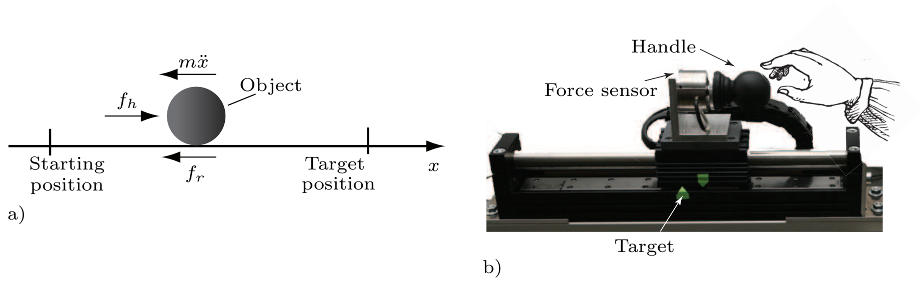

Fig. 2. Scenario and experimental setup: a) Object transportation and positioning; b) Admittance-type haptic interface

Experimental setup: Experiments are performed using a 1 DOF haptic interface controlled by position-based admittance control, see Fig. 2 b). The interface is equipped with a force, position, and acceleration sensor and uses a linear actuator Thrusttube 2504 of Copley Controls. Its control is implemented in Matlab/Simulink and executed on the Linux Real Time Application Interface

RTAI. The haptic rendering runs on another computer, and communication is realized by a UDP connection in a local area network.

Assistance function: In the experiments an assistance function is used which aims at compensating static friction to ease the positioning of the object. A possibility to realize such a compensation is achieved by amplifying the force f_h applied by the human when she/he tries to move the object:

$$f_v = \begin{cases} kf_h & \text{for } \dot{x} \leq \epsilon_{\dot{x}}, \\ 0 & \text{for } \dot{x} > \epsilon_{\dot{x}}; \end{cases}$$

where k is a predefined fixed gain. This assistance function virtually reduces the weight of the object and thus results into an easier and faster manipulation, see Fig. 3. During transportation, however, the assistance is switched off, which allows the human to feel the real mass and friction between object and environment.

Experimental results: The presented approach for intention recognition was evaluated by an experiment, where a subject who was familiar with the setup performed 10 trials (each two minutes long), consisting of random sequences of positioning and transportation tasks. The sequence of the tasks was determined by the subject and no special instructions were given how to concatenate them. The result of the intention recognition was visualized on a screen during runtime in form of text messages. Using a button the subject was asked to indicate whether the recognized intention was in agreement with her/his own intention. The used classifiers were trained offline by means of the Baum-Welch algorithm. To produce training data the subject performed each task 30 times in both, assisted and non-assisted mode.

In total 251 intentions were evaluated by the human, out of them 30.68% for positioning and 69.32% for transportation. The larger number of transportation tasks can be explained by the fact that transportation tasks consist of a single fast motion, while positioning tasks require several iterations even when assistance is applied and thus, their execution is more time consuming. Moreover, tasks were randomly performed following the subject preferences as indicated above. Please note that the human only rated the correctness of the identified intention while the effectiveness of the assistance function lied beyond the focus of the performed study.

In 70.73% of all cases, the recognized intention was in agreement with the feedback provided by the human. For the positioning task a recognition rate of 96.35% and for the transportation task a recognition rate of 59.6% was achieved.

As the probability of correct recognition by chance for our experiment is 50%, the overall recognition can be regarded as good. The lower recognition rate for transportation tasks can be explained by similar human behavior and thus observed patterns in both, positioning and transportation tasks, in case assistance is provided. This effect, however, can be reduced by dynamic adaptation of the symbol sequence length.

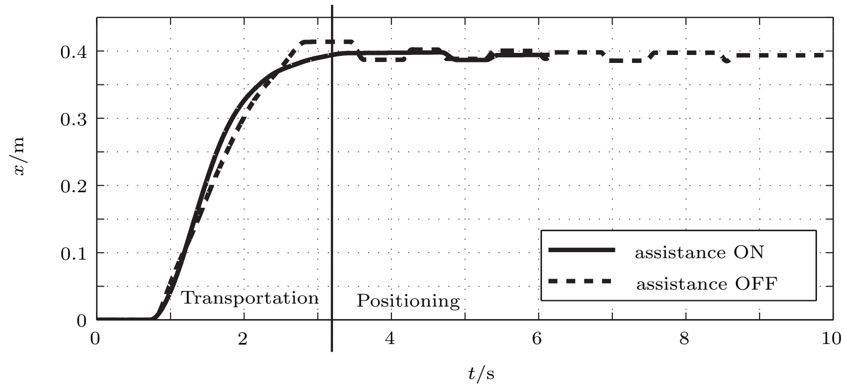

Fig. 3. Comparison of computer-assisted and direct manipulation in a transportation and positioning task

Discussion: The presented algorithm for intention recognition doesn't use any prior knowledge, like e.g. the sequence of tasks and thus, is most efficient when the behavior of the human is totally random. For such situations it clearly outperforms context-dependent methods that evaluate the second problem over HMMs to predict the next task to be executed.

Moreover, the proposed method is able to recognize human intentions, even in case assistance functions are activated. In such situations, intention recognition algorithms without compensation of the assistance function are doomed to failure, because of the changes in the observed patterns.

4 Conclusion and Future Work

In this work we presented an online intention recognition algorithm for computer-assisted teleoperation, which is able to reliable recognize the type of task the human is currently performing in case of assisted and non-assisted teleoperation. The approach is based on analyzing time domain haptic data recorded during execution of the task. The effect of the assistance function on human behavior is taken into account by training multiple classifiers for each type of assistance. The experimental results show good classification for stable human interaction behavior. Changing human behavior over time, however, reduces the recognition accuracy of the developed algorithm, and thus, a method for dynamic feature extraction to reduce this effect is currently under development.

So far the proposed approach requires an intensive training phase and provides reliable estimations only for the person who has provided the training data. Future work will consequently focus on reducing the required training data and fast adaptation of the classifiers to different human operators. Finally, the extension of the presented approach to tasks requiring more degrees of freedom is envisaged.

Acknowledgments

This work is supported in part by the German Research Foundation (DFG) within the collaborative research center SFB453 "High-Fidelity Telepresence and Teleaction".

References

1. B. Corteville, E. Aertbelien, H. Bruyninckx, J. D. Schutter, and H. V. Brussel, "Human-inspired robot assistant for fast point-to-point movements," in *2007 IEEE International Conference on Robotics and Automation*, pp. 3639–3644, IEEE, 2007.
2. U. Unterhinninghofen, F. K. Freyberger, and M. Buss, "Study on computer assistance for telepresent reaching movements," in *EuroHaptics '08: Proceedings of the 6th international conference on Haptics*, (Berlin, Heidelberg), pp. 745–754, Springer-Verlag, 2008.
3. D. Lee, D. Kulic, and Y. Nakamura, "Missing motion data recovery using factorial hidden markov models," in *Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on*, pp. 1722–1728, May 2008.
4. D. Aarno and D. Kragic, "Motion intention recognition in robot assisted applications," *Robot. Auton. Syst.*, vol. 56, no. 8, pp. 692–705, 2008.
5. R. Zollner, O. Rogalla, R. Dillmann, and M. Zollner, "Understanding users intention: programming fine manipulation tasks by demonstration," *Intelligent Robots and System, 2002. IEEE/RSJ International Conference on*, vol. 2, pp. 1114–1119 vol.2, 2002.
6. A. Castellani, D. Botturi, M. Bicego, and P. Fiorini, "Hybrid hmm/svm model for the analysis and segmentation of teleoperation tasks," in *Robotics and Automation, 2004. Proceedings. ICRA '04. 2004 IEEE International Conference on*, vol. 3, pp. 2918–2923 Vol.3, April-1 May 2004.
7. D. Aarno and D. Kragic, "Layered HMM for motion intention recognition," in *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on*, pp. 5130–5135, Oct. 2006.
8. N. Stefanov, A. Peer, and M. Buss, "Online intention recognition for computer-assisted teleoperation," in *IEEE International Conference on Robotics and Automation*, 2010, to appear.
9. R. Duda, P. Hart, and D. Stork, *Pattern Classification.* Wiley-Interscience Publication, 2000.