A Multimodal Communication Interface for Assistive Robots for the Elderly

Maria Javaid, Lin Chen, Uzair Ahmed, Miloš Žefran and Barbara Di Eugenio

I. Introduction

The RoboHelper project aims at building an effective and user friendly communication interface for assistive robots that can help the elderly live independently at home. Such communication interface should incorporate multiple modalities of communication, since collaborative task-oriented human-human communication is inherently multimodal.

Data was collected from twenty collaborative task-oriented human-human communication sessions between a helper and an elderly person in a realistic setting (fully functional studio apartment). We observed that these interactions involve a specific type of gestures, Haptic-Ostensive (H-O) actions [1]. H-O actions manipulate objects, but they also often perform a referring function. For example, the helper may open a drawer without saying a word while looking for a pot, and the elder may say *"Not there"*. *"there"* is a so-called *referring expression* whose *referent* (the drawer) was established in the context without any words, but through the action of opening the drawer itself, namely via visual and/or haptic cues (note that as far as the helper is concerned, even if they were blindfolded they could establish the existence of the drawer via the action they perform). We ultimately envision an interface that can exploit a multiplicity of cues – spoken, visual and haptic – to support communication. In our project, we investigate which roles H-O actions play in interaction, since H-O actions have not been studied as much as spoken and visual signals; and we show that we can actually recognize H-O actions from the haptic signals. We believe that our work on H-O actions is applicable to any domain where participants (including robots) collaborate on tasks that involve object manipulation.

A. Data Collection and Analysis

Speech, vision and haptics data was collected while an elderly person was helped by a caregiver to perform (Instrumental) Activities of Daily Living (ADLs and IADLs) [2] in a fully-functional studio apartment. ADLs are activities that are essential for a person to live independently, such as getting up from a bed or chair, getting dressed, preparing dinner. Our experiments were designed in such a way that the elderly person was in charge of the ADLs with the caregiver providing only minimal support, since gerontological studies indicate that physical and cognitive activities keep older people healthier [3], [4]. The haptic data was unobtrusively obtained in our studies via a sensory glove with pressure sensors that we developed in-house [5]. This data constitutes the ELDERLY-AT-HOME corpus which was extensively annotated for features of interest (details in [6]). This corpus comprises 2555 spoken utterances, 571 pointing gestures, and 1088 H-O actions - hence, H-O actions constitute 25% of all communication events.

II. H-O Actions: Roles and Recognition

A. H-O Actions and Language Understanding

To process the spoken utterances that we collected in RoboHelper, we developed two components:

1) The Dialogue Act Classifier, which recognizes the true intention behind an utterance. For example, *Not there* is not simply a negative statement, but an indirect command to look somewhere else.
2) The Reference Resolution Module, which resolves referring expressions to the entities that they refer to in the real world (referents). Examples of referring expressions from our corpus include pronouns such as *it, they*; deictics such as *this, that, there*; and full noun phrases such as *the second drawer*.

For both components, we showed that the information conveyed by H-O actions plays a crucial role – namely, that including H-O actions (and gestures in general) among the features that these two modules have at their disposal significantly imrpoves their performance [6], [7].

B. Recognition of H-O Actions from Haptic Signals

The components we just described make crucial use of H-O actions, however those models are based on H-O actions *manually annotated* on the basis of the videos we collected. Needless to say, to achieve true human-robot interaction, H-O actions need to be recognized from the haptic signals. Ideally, we should automatically recognize the H-O actions from the haptic data collected in our corpus. Unfortunately, this was not possible because the collected haptic data turned out to be corrupt. In the *a posteriori* analysis, several problems were identified. In some cases, the communication was lost between the data collection module and the computer used to store the data (the speech and video data was not affected). In other cases, the pressure sensor data was corrupted due to an outer glove that was worn over the fabricated glove so that the subjects' hands would be more easily identifiable by the computer vision algorithms. Thus additional data

This research was supported in part by NSF grants IIS-0905593, CNS-0910988 and CNS-1035914.

M. Javaid, U. Ahmed and M. Žefran are with the Dept. of Electrical and Computer Engineering, while Lin Chen and B. Di Eugenio are with the Dept. of Computer Science, U. of Illinois at Chicago, Chicago, USA. Email: {mjavai2, lchen43, uahmed9}@uic.edu, {mzefran, bdieugen}@uic.edu

was collected in the same studio apartment to develop the automatic recognition algorithms for H-O actions. Four pairs of subjects wore the same equipment as in the ELDERLY-AT-HOME data collection, and performed the same (I)ADLs. Our subjects were young adults (UIC students); one subject played the role of the helper, and the other the role of the elderly person. To collect the haptic data the subject playing the role of helper wore our data glove. The data obtained from the experiments mirroring the ELDERLY-AT-HOME corpus was termed as the Naturalistic Validation (NatVal) set. Table I presents the frequency distribution of H-O actions among the four helper subjects, in the *NatVal* set. Table II shows the recognition results for H-O actions performed in the *NatVal* set. Overall, 67.9% of the total 183 H-O actions were correctly classified to the right group. Elsewhere we have provided further details on the successful classification of different H-O actions within one group (manipulation of planar object) [5]. This classification is performed using the Dynamic Time Warping (DTW) algorithm. Results obtained with other algorithms have been discussed in [8].

TABLE I: Frequency distribution of actions in the *NatVal* set

	Helper 1	**Helper 2**	**Helper 3**	**Helper 4**	**Total**
Open/Close Cabinet	13	13	24	17	**67**
Open/Close Drawer	16	0	17	4	**37**
Grasp Plate	3	4	2	8	**17**
Grasp Pot	8	6	5	4	**23**
Grasp Small Items	3	6	4	4	**17**
Idle Hand	8	4	8	2	**22**
Total	**51**	**33**	**60**	**39**	**183**

TABLE II: Confusion Matrix for Cross-Validation on the *NatVal* set

	Open/ Close Cabinet	**Open/ Close Drawer**	**Grasp Plate**	**Grasp Pot**	**Grasp Small Items**	**Idle Hand**	**Total**
Open/Close Cabinet	**53**	5	3	1	0	5	**67**
Open/Close Drawer	2	**28**	0	0	3	4	**37**
Grasp Plate	4	2	**6**	0	3	2	**17**
Grasp Pot	4	1	1	**17**	0	0	**23**
Grasp Small Items	5	2	2	2	**5**	1	**17**
Idle Hand	1	4	1	0	1	**13**	**20**
Total	**69**	**42**	**13**	**20**	**12**	**25**	**183**

C. Interpersonal Communication Through Physical Interaction

In addition to recognizing H-O actions, the research team identified actions which involve interpersonal communication through haptics and explored one such action, i.e., collaborative manipulation of planar object. The details of this work are presented in [5]. This work can be easily adapted to different sensors and hardware platforms and particularly, utilized to make the robot-human hand-over of planar objects more realistic.

III. Future Work

The research done as a part of the RoboHelper project is an important step towards better communication interfaces for assistive robots, in general and specifically in the elderly care domain. Future work includes testing the developed methodology on a robotic platform. A preliminary implementation is underway in ROS [9], including a real-time implementation of the H-O action recognition algorithms [10]. We have started experimenting with a Nao robot - a humanoid robot from Aldebaran Robotics. An important question that will be addressed is whether H-O actions can inform coreference resolution and dialogue act classification even when they are recognized automatically and thus with lower accuracy.

References

[1] M. Foster, E. Bard, M. Guhe, R. Hill, J. Oberlander, and A. Knoll, "The roles of haptic-ostensive referring expressions in cooperative, task-based human-robot dialogue," in *Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction*. ACM, 2008, pp. 295–302.

[2] K. M. Krapp, *The Gale Encyclopedia of Nursing & Allied Health*. Gale Group, Inc. Chapter Activities of Daily Living Evaluation., 2002.

[3] M. Angevaren, G. Aufdemkampe, H. Verhaar, A. Aleman, and L. Vanhees, "Physical activity and enhanced fitness to improve cognitive function in older people without known cognitive impairment," *Cochrane Database Syst Rev*, vol. 3, no. 3, 2008.

[4] E. Kraft, "Cognitive function, physical activity, and aging: possible biological links and implications for multimodal interventions," *Aging, Neuropsychology, and Cognition*, vol. 19, no. 1-2, pp. 248–263, 2012.

[5] M. Javaid, M. Žefran, and B. Di Eugenio, "Communication through physical interaction: A study of human collaborative manipulation of a planar object," in *Proceedings of the 23rd IEEE International Symposium on Robot and Human Interactive Communication*, 2014.

[6] L. Chen and B. Di Eugenio, "Multimodality and Dialogue Act Classification in the RoboHelper Project," in *SIGDIAL*, Metz, France, August 2013, pp. 183–192.

[7] L. Chen and B. Di Eugenio, "Co-reference via Pointing and Haptics in Multi-Modal Dialogues," in *Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Montréal, Canada, June 2012, pp. 523–527.

[8] M. Javaid, M. Žefran, and B. Di Eugenio, "Using pressure sensors to identify user actions during physical interaction," submitted to ASME 2014 Dynamic Systems and Control Conference.

[9] M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, and A. Y. Ng, "ROS: an open-source robot operating system," in *ICRA Workshop on Open Source Software*, 2009.

[10] U. Ahmed, "Implementing the Robot Modules in ROS," Master's thesis, University of Illinois Chicago, 2014.