The MOBOT Rollator Human-Robot Interaction Model and User Evaluation Process

Eleni Efthimiou, Stavroula-Evita Fotinea, Theodore Goulas, Maria Koutsombogera, Panagiotis Karioris, Anna Vacalopoulou
Institute for Language and Speech Processing/"Athena" RC
Athens, Greece

Isidoros Rodomagoulakis, Petros Maragos, Costas Tzafestas, Vassilis Pitsikalis
Institute of Communication and Computer Systems–NTUA,
Athens, Greece

Yiannis Koumpouros
Technological Educational Institute of Athens, Department of Informatics
Aigaleo, Greece

Alexandra Karavasili, Panagiotis Siavelis, Foteini Koureta, Despoina Alexopoulou
Diaplasis Rehabilitation Center
Kalamata, Greece

***Abstract*—In this paper we discuss the integration of a communication model in the MOBOT assistive robotic platform and its evaluation by target users. The MOBOT platform envisions the development of cognitive robotic assistant prototypes that act proactively, adaptively and interactively with respect to elderly humans with slight walking and cognitive impairments. The respective multimodal action recognition system has been developed to monitor, analyze and predict user actions with a high level of accuracy and detail. The robotic platform incorporates a human-robot communication model that has been defined with semantics of human actions in interaction, their capture and their representation in terms of behavioral patterns, to achieve an effective, natural interaction, aiming to support elderly users of slight walking and cognitive inability. The platform has been evaluated in a series of validation experiments with end users, the procedure and results of which are also presented in this paper.**

Keywords—multisensory data, multimodal HRI model, multimodal semantics, multimodal human-robot communication, natural HRI, evaluation, assessment

I. Introduction

The need to improve the quality of daily living by supporting mobility and vitality in our ageing society, as well as enhance independent living of elderly individuals with motor limitations [1] has inspired technological solutions towards developing intelligent active mobility assistance robots for indoor environments, providing user-centered, context-adaptive and natural support [2-5]. Cognitive and robotic architectures may be able to provide advanced interactive capabilities with humans and influence the usability and functionality of a resulting system, contributing to its quality of services. Such systems may not only integrate multiple advanced cognitive abilities, but also employ methods that are extendible to various other robotic and non-robotic applications required in assisting humans with mobility disabilities. The MOBOT project[1] has addressed this need envisioning cognitive robotic assistants that act (a) proactively by realizing an autonomous and context-specific monitoring of human activities and by subsequently reasoning on meaningful user behavioral patterns, as well as (b) adaptively and interactively, by analyzing multi-sensory and physiological signals related to gait and postural stability, and by performing adaptive compliance control for optimal physical support and active fall prevention.

To address these targets, a multimodal action recognition system has been developed to monitor, analyze and predict user actions with a high level of accuracy and detail. Parallel to the enhancement of computer vision techniques with modalities such as range sensor images, haptic information and command-level speech and gesture recognition, data-driven multimodal human behavior analysis has been conducted in order to extract behavioral patterns of elderly people. The aim was to import the basic elements of these behavioral patterns into a multimodal human-robot communication system [6], involving both verbal and nonverbal communication conceptually and systemically synthesized into mobility assistance models after taking into consideration safety critical requirements.

Near the end of the project, the different modules have been incorporated in a behavior-based and context-aware robot control framework aiming at providing situation-adapted optimal assistance to users [7]. Direct involvement of end-user groups in various stages of the development of prototypes has ensured that actual user needs are addressed by the functionalities and communication capabilities of the platform's prototypes. Thus, user trials have been conducted to evaluate and benchmark the overall system.

[1] www.mobot-project.eu/

In the following sections, we report on the technologies integrated in the robotic platform's prototypes and the respective functionalities they provide (Section II), the human-robot interaction (HRI) communication model adopted (Section III), and the different end user evaluation and usability studies (Section IV). Emphasis is placed on the final platform evaluation, which tested both the integrated technologies and the adopted communication model.

II. Technologies and Functionalities Integrated in the MOBOT Platform

Technologies integrated in the MOBOT platform had to reach a significant level of enhancement and achieve the state of maturity required to meet the set targets in respect to functionalities and safety controls envisioned. In this respect, the integration and synergies of a wide range of technologies have made the development of the MOBOT platform a rather ambitious endeavor. In this section, we briefly presented the various types of integrated technologies and the respective functionalities they support, to illustrate how the adopted multimodal HRI communication model makes optimal use of the available technological solutions. The MOBOT multimodal-multisensorial dataset [8] has been exploited towards the enhancement of all technologies explored in order to be integrated in the MOBOT robotic platform.

A. Platform Technologies

The platform capability to detect human activity and drive subsequently user intentions e.g. robot activation [9-10], was supported by Research and Development work on visual action recognition in video streams captured by visual sensors mounted on the MOBOT robotic platform [11], on object detection and on advances in human body pose estimation. The MOBOT dataset has been exploited along with previous research work on action/gesture recognition, and provided the workbench of newly developed algorithms. In this respect, advancements comprise the development of an improved gesture recognition method based on specific articulatory points such as the arms and the hands of the subject as well as the application and experimentation on actual MOBOT data, following research work with other datasets.

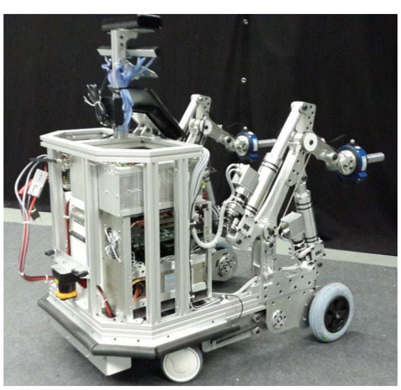

Fig. 1. The active rollator used during the evaluation of the MOBOT rollator-type mobility assistant.

HMM-type classifiers, using as feature extraction visual cues the handshape (provided by the RGB stream) and 3D movement-position (provided by Kinect's depth stream and skeleton tracking) [12-13] formed the experimental framework [14].

Moreover, a spoken command recognition system has been employed; which uses a voice activity detector for time segments identification of the commands in the audio stream and next exploits an HMM model set on these segments [14]. Experimentation was initially tested on the audio data from the benchmark dataset of the ACM 2013 Multimodal Gesture Challenge [15] and also on MOBOT data during the development of the complete MOBOT spoken command recognition system, also integrated on the ROS platform with a MEMS microphone array mounted on the MOBOT active rollator prototype [16].

As regards multimodal sensor fusion for audio-visual gesture recognition we exploited color, depth and audio information captured by the Kinect sensor, recognized time sequences of audio-visual gesture commands based on fusion of all different cues and modalities and generalized using an activity detection component. The adopted approach provided results which greatly outperform all other competing published approaches on the ACM 2013 benchmark dataset and achieve a 93% accuracy, which corresponds to a 47% error reduction over the best competing approach. Finally, haptic data processing exploited data of the two force/torque sensors mounted on the handles of the prototype. The sensors were used to detect and quantify haptic HR interactions. Typical interaction patterns that have been identified include standing up, sitting down and walking actions [17].

Furthermore, the processing of physiological data focuses on user fatigue, which is considered a crucial physiological state that can strongly affect human performance. Fatigue estimation is based on two specific features, i.e. the human heart rate and the total work performed. Moreover, available fatigue indicators suitable for elderly fatigue estimation [18] are extended in order to be used in the context of mobility assistive robots.

B. Platform Functionalities

The platform consists of two robots, a rollator-type robot for walking and sit-to-stand (STS) assistance and a nurse-type robot for sit-to-stand assistance. The MOBOT rollator is an assistive device which includes the main frame, the actuated handles, the active wheels, a user interface, an electronic control unit and various environment and user sensors [14].

The distinct functionalities meet both mechanical design and requirements of the users. The functionalities of the MOBOT rollator are classified in the following seven categories, each of which demonstrates the capacity of the rollator in terms of the following [14]:

1) Perceiving the user and adapting itself to the user. This includes: localization of the user in connection with the rollator through the use of 3D coordinates and the distance between user and the variables "distant", "close" and "in contact"; tracking of the articulated human body; detection of walking patterns; recognition of user gestures, actions, plans, and intentions; recognition and interpretation of the voice commands; monitoring human performance and postural stability as well as detecting unstable configuration or falls; and, finally, recognition of user physiological state.

2) Detecting the environment, such as the existence of obstacles, locomotion related data e.g. type/slip of surface, environment related data e.g. slopes, and creating a map of the surroundings.

3) Localizing itself within the aforementioned map.

4) Approaching the user from a distance.

5) Assisting the user by: offering physical assistance in sit-to-stand and stand-to-sit transfers; offering assistance to the user in three ways when walking, including following the intention of the user ("dock" to the user, accelerate, maneuver, decelerate, stop); balancing and stabilizing the user (including fall prevention), and offering assistance when moving through narrow passages and opening/closing doors; providing assistance to the user when standing (in proximity but in no contact mode); following him/her; offering sensorial help, in order, for instance, to avoid obstacles –be it static or dynamic, positive or negative– or to cope with slopes; offering cognitive assistance; and, finally, assisting the localization of the user by guiding/navigating him/her.

6) Leaving the user and assume the parking position in autonomous mode.

7) Charging autonomously.

The aforementioned attributes are connected to the technological solutions that were mentioned in section II.A and have served as the bases for the multimodal communication model of the platform.

III. The MOBOT Multimodal HRI Communication Model

The synergies among the MOBOT modules have been facilitated through integration of related technologies, so that the platform achieves to offer walking and cognitive assistance to users who are elderly and have minor walking and cognitive problems. In this respect, the compass towards implementing the capabilities of the platform has been a well-articulated HRI communication model which enhances the communication between the platform and its user in the most natural possible manner. In doing so, the model is based both in the degree of the incorporated technologies and the ways through which the target group of users normally communicate [14].

The multimodal dataset [8] which was attained in the initial phases of the project and the relevant annotation of both verbal and gestural communicative signals were based on the basis of defining a multimodal human-robot-interaction model (HRI) which would resemble natural human behavior as much as possible. The dataset was acquired through the use of a sensorized passive rollator including multimodal input from several sensors, microphones, and cameras, as well as a motion capture system employed to record human limb movements and the rollator and subject's absolute positions in space. The MOBOT corpus is one of the rare samples of multimodal-multisensorial resources with a fairly rich content of activities carried out by humans in interaction with a device. The aforementioned communication model was put together as a structured tree of possible multimodal interactions of the action-reaction type, involving audio and gestural signals which were enriched by a number of cognitive assistance assertions from the part of the platform, assimilating human reinforcement to elderly individuals when performing an everyday activity [14].

Interaction on the part of the platform is possible in three modes: i) hands-on, ii) following, and iii) stand-by. At the same time, user input is used to develop a multimodal dialogue strategy, which allows for two communication options. The first option is body posture in silence, which means total absence of any signal via speech or gesture and processing information which may be connected to the action recognition module of the platform. The second option is the existence of speech and/or gesture signals, processing information which may be connected to the audio and gestural signal recognition (Fig. 2). The capacity of the system to navigate through the use of a map is connected to a number of cognitive support messages. These messages reach the user as oral questions, which resemble the ones asked by human carers in human interaction. These messages connect to the activation and approach of the robotic device, the process of decision making when it comes to choosing a route or to avoiding obstacles. Some kind of verification by the user is required upon hearing these messages, which encourage users to complete activities successfully [14].

Fig. 2. The MOBOT active rollator in audio-gestural interaction with an elderly user.

Actual user needs served as the basis on which the use scenarios were developed. These needs had been identified from both the MOBOT dataset and the intermediate evaluation of end users. This was because the primary concept behind acquiring the data is that, for a given system to try and model human actions based on interactional behavior, it is essential to know the structures of human actions. More to the point, assistive systems relating to human-machine interaction need to be in position to analyze human behavior into measurable features which are also machine detectable. This is what helps them select and plan appropriate support actions based on heterogeneous sensory data [6; 14; 19].

In this particular case, elderly people of varying ages, genders, motor and cognitive capacities performed a variety of activities for which assistance was needed. The scenarios recorded involved human carers and a passive rollator[2] and were based on a list of use cases comprising of actions that typical end users will typically need to perform in realistic circumstances [14].

IV. EVALUATION AND VALIDATION OF THE MOBOT PLATFORM BY END USERS

When developing assistive devices, user evaluation and acceptance is considered one of the major factors in determining the success rate of the whole process; from device overall design to integration of foreseen technologies and implementation of interface with the user. Along these lines, end user evaluation and validation has been one of the most significant targets of the MOBOT platform implementation. To this end, the platform functionalities have undergone steady evaluation cycles in several milestone development phases in order to be informed as to overall user needs coverage in respect to the initial MOBOT targets. All trials involved the MOBOT rollator-type mobility assistant.

A. *First Evaluation Phase*

The first evaluation phase took place at the Geriatric Centre of BETHANIEN Hospital of the University of Heidelberg between late October and early December 2014. 36 patients, who were selected according to well-defined criteria, formed the participants group of the first evaluation study. The members of this group were recruited from different environments supervised by the BETHANIEN-Hospital, such as the institution's geriatric rehab wards, as well as various nursing homes and a rehab sports club. This evaluation study involved the first edition of the project's active rollator and aimed at validating the specific functionalities which were mounted on the device at the specific time point. The functionalities undergoing user evaluation included the implemented sit-to-stand assistance (STS), the obstacle avoidance functionality testing the device's sensorial assistance ability, and also the platform's basic walking assistance [14]. In order to best exploit the results of this critical first evaluation, the design of the process which actually followed focused on obtaining valid information from the part of end users. Thus, functionality testing was based on the execution of a set of scenarios, for which performance was assessed by means of methods specifically designed to incorporate both qualitative and quantitative measures. A gain of this trial was that the implemented assessment scheme proved to be successful in respect to addressing the user group needs and capabilities. This first experience also fed the trials, evaluation studies, and validation procedures which took place next, throughout the project lifecycle.

As regards the assistive functionalities tested, the first evaluation results demonstrated that all of them were in the right direction regarding end user needs support. Moreover, the feedback received in respect to the subjective user perception of the MOBOT device was also positive. A detailed account of the MOBOT platform first evaluation study, including a thorough report on: i) the procedures followed to form the testers group, ii) the scenarios implemented to test the functionalities mounted on the rollator device, iii) assessment metrics, and iv) a summary of the obtained preliminary results, is available in [20].

To complete the picture, an extensive survey of available devices and their related validation studies details are provided in [21] along with a report on the use cases selected for the MOBOT trials, while an updated version of this information can be found in [22].

Typically, use cases reflect real life scenarios of use, which allow to further enhance technical development and validate a device in the process of manufacturing. Within the MOBOT project framework, use cases needed to be defined, taking into account the specific setting of geriatric rehabilitation and long-term care. The adopted use cases are composed of (sequences of) task performance characterized by different duration, frequency of activity and clinical relevance, and executed by individuals of the specific user profile in various settings of their everyday life. Thus, some use cases need to take only seconds to perform, while others require several minutes. Similarly, a subset of activity environments appears quite often during a day, where others may set up very rarely. However, although some of the situations included in the testings may occur pretty rarely, they are crucial as regards the requirement on the part of the device to support the motor stability of the user, and in this way restrict the danger of falls. The three criteria of accuracy, validity, and reliability were assessed as regards the platform's technical functionalities. Clinical evaluation, on the other hand, prioritized the interaction between a human and the rollator [14].

In the first evaluation phase, three different user modes were tested by elderly users concerning these modes during the MOBOT HRI task: 1) Audio, 2) Gestural, 3) Audio-gestural. For each mode, a sequence of five commands was repeated three times in German, i.e. the users' native language. Participants were guided with prompts in order to perform the commands, seated on a chair 1-2 m behind the MOBOT rollator-type. Based on the data that has up to now been processed, it should be noted that there is a short percentage (approx. 11-15%) of missing scenario cases of data due to several reasons, e.g. the subject did not manage to conduct the experiment. More specifically, the patients who participated in the audio, gestural, and audio-gestural experiments were 18, 19 and 20, respectively. Based on these data, the obtained recognition accuracies for the single-modality experiments were 64.79%, for audio, and 37.87% for gestures, averaged across participants. Fig. 3 presents accuracies for the multimodal audio-gestural experiment, the obtained mean performances of the different modalities being: 52.64% for audio, 30.29% for gestures, and 58.30% for the audio-gestural mode. The results are promising based on the fact that the employed system was performing close to real-time on elderly users. Note that the proposed fusion scheme yielded 12% relative improvement against the best modality. Additionally, there was a 6% absolute improvement after ignoring false-alarm penalties on the computation of accuracy scores.

2 This process took place in the rehabilitation centre Agaplesion Bethanien Hospital/Geriatric Centre at the University of Heidelberg.

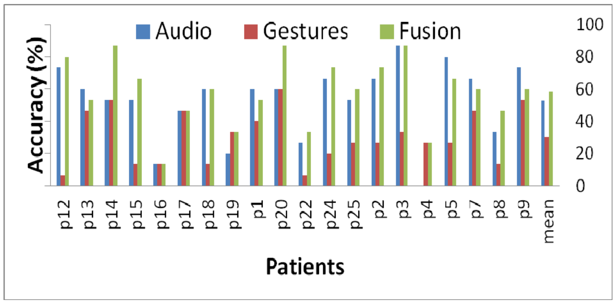

Fig. 3. Online command recognition accuracies on the audio-gestural scenario using audio-only, gestures-only and multimodal recognition.

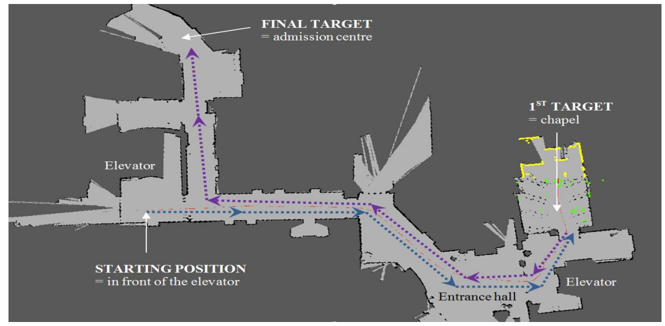

Fig. 4. Layout of the navigation trail at the ground floor of the BETHANIEN hospital.

B. Final Evaluation Phase

The final evaluation of the MOBOT prototypes was conducted in two phases: from April 22 to May 12 2016 at the BETHANIEN-Hospital, and in the period July 15-27 at the DIAPLASIS rehabilitation center in Kalamata, Greece, involving more than 30 subjects, complementing the sample of the previous evaluation cycles. A detailed account of the final user evaluation phase as implemented in the two validation sites and incorporating activities, hosted at the project's clinical partners' wards, can be found in the project deliverable D5.4 [23], which also provides an extensive account of validation procedures and results as regards the used prototype.

The aim of this evaluation study was to validate the cognitive (navigation) assistance system and the audio-gestural human-robot communication. According to these functionalities, two specific test scenarios with tailored assessment methods and qualitative as well as quantitative performance metrics were developed. Major parts of performance metrics are based on technical data derived from the data flow of integrated systems on the MOBOT rollator-type mobility assistant.

The first scenario implemented in this validation study aimed at assessing the performance of an audial cognitive assistance mode of operation. This mode is supported by a subset of the autonomous robot navigation modules that have been previously developed and successfully integrated on the MOBOT rollator platform, including mapping, localization and path planning (Fig. 4). This audial cognitive assistance module operates by providing prespecified audio cues to the user (Fig. 5), essentially associated to navigation instructions, while walking with the rollator. The module assumes a previously captured map of the indoor environment along with a set of preset guard points (2D positions) with associated audio tokens. The cognitive assistance module has also been augmented to accommodate loops and overlapping guard points within higher-level paths integrating intermediate target points. This module is designed to help users move around and orient them in an indoor environment by providing localization information as well as by giving assistive directions and issuing navigation instructions in the form of audial messages. The research hypothesis tested by this scenario is related to the effect that such a navigation assistance modality can have with respect to the cognitive status of the user (that is, for cognitively impaired vs. non-impaired subjects).

During the experiments, the MOBOT rollator-type device was assessed using both objective (success rate; task completion time; stopping time; walking trajectories; gait parameters captured by laser sensor) and subjective (ATD PA Greek version; QUEST 2.0 Greek version; PYTHEIA questionnaire) performance measures. In order to measure subjective user satisfaction, extensive research was conducted as regards available valid and reliable questionnaire types that could be used in the frames of the MOBOT experiments. As mentioned in [33], most of the studies are utilizing either custom-made questionnaires or interviews that are neither valid nor reliable instruments to represent the subjective opinion and perception of the end users. A new questionnaire was, therefore, developed to meet the needs of our study and to be available for use in other studies working in the development of rehabilitation or assistive robot devices [34, 35].

The second scenario tested a set of audio-gestural human-robot interactions, aiming to evaluate the mechanisms implemented and build a close to natural communication. Recent technological and scientific progress in assistive living and computer vision has led to significant interdisciplinary advances in assistive human-robot interfaces (for a review, see [14; 24-25; 29-31]). HR interfaces can offer natural communication channels to elderly subjects placing effort on making their interaction easier and on increasing the multimodal communication throughout.

This scenario aimed at validating the MOBOT action recognition system and models with real elderly subjects. This task still presents several challenges, as shown by the sparse similar works in the recent interdisciplinary literature. Visual and multimodal audio gestural action detection and recognition include several issues that can be considered challenging. For instance, action/gesture recognition independent of the subject's state and despite recent progress [32] has still a lot to gain from incorporation of rich temporal information and core machine learning enhancements due to the lack of adequate available data on current datasets.

The analysis and post-processing of results for the final evaluation studies are still ongoing. This is due to the large quantities of data that have been acquired. The results of these experiments and studies are to be published subsequently. Until then, we next present the experimental setup of the second phase of the final evaluation study (Kalamata, Greece) as well as some preliminary results and discussion.

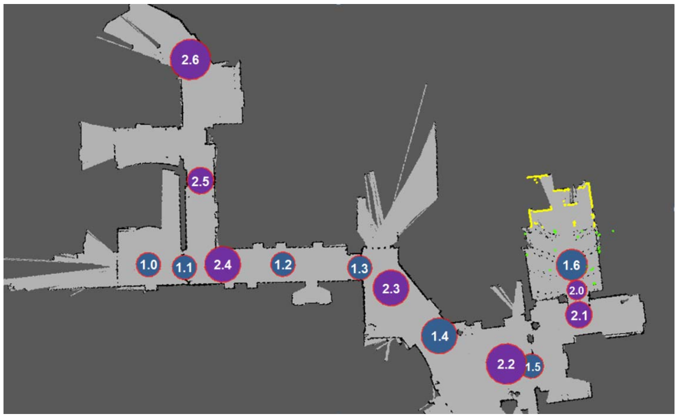

Fig. 5. Position of the audio cues provided for cognitive assistance through the navigation trail at the ground floor of the BETHANIEN hospital.

Three different user modes of communication between the participants and the MOBOT rollator-type mobility assistance were tested:

1. Audio mode
2. Gestural mode
3. Audio-gestural mode

For each mode, the following five commands were tested:

1. *Come near me!*
2. *I want to stand up*.
3. *I need help.*
4. *Where am I?*
5. *Go park!*

All participants were seated on a chair 1-2m behind the MOBOT rollator-type mobility assistant. Throughout the final evaluation studies the same execution protocol was followed.

The implementation of this scenario was based on the assumption that the subjects would utter and perform the audio-gestural commands exactly as they appeared in the prompts. However, this proved not to be the case at all during task performance. Thus, due to such parameters, initial results on performance metrics may be heavily revised after detailed manual annotation of the ground truths and the acquired multimodal data.

In this context, for the extraction of performance metrics H is the number of correct words, I the number of insertions (i.e. words present in the recognizer's transcription but not in the reference), and N the total number of words in the reference transcription. The percentage of correctly recognized words is given by Correct = H/N × 100%, and the word recognition accuracy is computed by Accuracy = (H – I)/N × 100%.

On the basis of the limited data currently being processed, the first thing to be spotted is a 11-15% of missing parts in the experimental scenarios, which is due to several reasons, such as the subjects' inability to complete the entire scenario.

The median automatic recognition rate in terms of accuracy has up to now been reported as 53.3% for the audio and 40% for the visual modality. On the other hand, there are subjects performing as low as 6 or 26% in either the audio or visual modalities. However, the entire set of data should be carefully reviewed not only automatically but also by manual annotation, so as to refine the ground-truth labels and find out how "true" these results are in reality.

The assessment of user satisfaction from their experience of using robotic assistive technology is a rather difficult process. Thus far, no standard methodology or scale has been reported. Researchers tend to use custom-made questionnaires in order to investigate user satisfaction. This, however, drives in non-validated scales/questionnaires and makes it difficult to compare the results from different researchers. A systematic survey on the existing scales that could be used was conducted recently and is reported in [33]. Since there was no prior scale that could satisfy the requirements of the subjective evaluation of MOBOT, a new scale was developed and tested in terms of its validity and reliability [34-35]. Participants in the experimentation/evaluation were asked to fill in several assessment questionnaires (i.e. PYTHEIA, QUEST 2.0 and ATDPA) [33-37], adapted to MOBOT evaluation criteria. The assessment questionnaires measure user experience and satisfaction with the implemented functionalities. The newly developed PYTHEIA questionnaire consists of 15 main items incorporated in Part A of the questionnaire, and 5 more items in Part B, that can be applied to assess user satisfaction not only with the device as a whole, but also with any individual functionality integrated in the evaluated platform. This allowed the evaluation of the cognitive assistance module and the audio-gestural command module of the MOBOT rollator independently. Indicatively, some of the characteristics assessed using the aforementioned questionnaires are ease of use, performance expectancy, reciprocity, feeling of security, usability, adaptability, societal impact etc.

Finally, in order to test the navigation and communication models presented previously, a new and more complex navigation trail was designed at DIAPLASIS. Patients with different pathologies and characteristics (e.g. multiple sclerosis, myopathies, Parkinson's disease) tested MOBOT under real conditions.

According to the results obtained from the validation phase at DIAPLASIS, only three subjects failed to complete the cognitive assistance experiment. The subjective evaluation of the MOBOT platform using the aforementioned three questionnaires (i.e. QUEST 2.0, ATDPA, and PYTHEIA) showed that the users of the MOBOT rollator were generally very satisfied with the implemented functionalities, while the human-robot communication model was found satisfactory and helpful. More specifically, the subjects scored the MOBOT with mean=3.91 in the PYTHEIA scale, mean=3.82 in the ATDPA and mean=3.71 in the QUEST 2.0 scale. The Likert scales used in all questionnaires had 5 as the highest score option (PYTHEIA-5=Extremely satisfied, ATDPA-5=Satisfied all the time, QUEST-5=Very satisfied). While using the PYTHEIA Part-B questionnaire, the subjects assessed the cognitive assistance functionality with 4.46 and the audio-gestural functionality with 4.30, indicating their high level of satisfaction from the use of these functionalities. Moreover, they ranked the three most important features of the MOBOT rollator as follows: 1st) ease of use (median=4, mean=4.2759, SD=0.84077), 2nd) felling of security (median=4, mean=4.2759, SD=0.84077), 3rd) contribution to everyday life improvement (median=4, mean=3.7931, SD=1.04810). During trials, this judgment was expressed with exclamation remarks at the rollator's response to the HRI commands but also in respect to navigation guidelines and localization information provided by the rollator during actual navigation.

V. Discussion

This work reported on the final evaluation of the MOBOT robotic platform based on the integration of a related technologies and a multimodal communication model to support a natural multimodal HRI. It presented the various types of integrated technologies and the respective functionalities they support, to illustrate how the adopted multimodal HRI communication model makes optimal use of the available technological solutions. The communication model was built through the analysis and annotation of acquired multi-sensory data of users interacting with mobility aids and their emerging multimodal human behavioral patterns. In parallel, the analysis of walking patterns and critical safety requirements in mobility assistance, as well as the detection of abnormalities were taken into account for the model development. In parallel, a goal that has been addressed is the implementation of context-dependent instantiations and flows of the communication model following the different interaction states and user needs successfully through the use of the audio and visual channels of communication with the device.

The adaptation of anticipated human behavior in the artificial cognitive system promotes the proactive assistance and the contextual mobility support to humans. Thus, the active involvement of end-users together with iterative trials and evaluation of the artificial cognitive system provided continuous consultations and user feedback, and thus guides research and development towards real use cases. Since the aim of the platform is to provide situation-adapted optimal assistance to users, the platform functionalities and communication capabilities have been tested by end user groups to evaluate and benchmark the overall system.

The initial trials were positively assessed in terms of feasibility and adequacy to the user needs (especially regarding the basic walking assistance, and the subjective user perception of the MOBOT device and its assistance systems), and it has served as a valuable experiment for planning all following trials and evaluation studies.

The final evaluation phase performed at DIAPLASIS, included a relatively extensive number of end users towards an overall benchmark verification of the developed intelligent mobility assistants. A total of 30 subjects participated in this evaluation study completing both scenarios of the experiments. A lengthy recruitment process had already taken place in order to satisfy the needs of the evaluation. Since the evaluation was conducted during the last days of the MOBOT project and the collected data are still being processed, it is not possible to report any statistically processed results at this stage. However, it is noteworthy to mention that the diversity of the participants in terms of pathology, age and health condition can be considered as an asset for the evaluation of MOBOT under real conditions.

According to the early data collected during this final evaluation phase, the patients who used the MOBOT rollator reported satisfactory results regarding the MOBOT as a whole, the communication model and the navigation feature. Even though the average ranking was satisfactory, some difficulties were faced by users with serious problems in speech articulation when using oral commands to guide the MOBOT platform. The gesture mode was problematic only to those with very limited ability to move their upper arms.

Similarly, continuous repetition of localization information by the rollator proved to be a bit annoying for users with no mental impairment, while users with various degrees of mental disorders exhibited full satisfaction of the support they received and the information repetition tempo, which suggests that this feature should allow a degree of adjustability according to specific user needs.

VI. Conclusions

To summarize, the preliminary results on objective performance metrics of the cognitive assistance scenario look promising, with indications that cognitively impaired participants achieved a better user performance (i.e. task completion time) with the assistance of the MOBOT's cognitive system when completing a navigation task in a complex real-world application scenario. Results on the subjective user satisfaction also look positive, suggesting that the use of the MOBOT functionality does not cause users any feelings of insecurity, is interesting and challenging to them, and may have a positive effect on their life quality. Although preliminary, these results present a clear indication of the potential added value of such a robotic functionality for cognitively impaired, elderly persons.

As audio-gestural human-robot interaction enables new modes of natural communication between users and assistive robotic devices, it becomes particularly relevant for the development of such devices focusing on users. Preliminary results on the automatic recognition accuracy of the MOBOT's multimodal action recognition system also reflect this issue, pinpointing the need for further research. Furthermore, results of the audio-gestural scenario, where user performance is more emphasized, indicate that in order for human-robot communication to be adequately successful, further focus should be placed on the training of the user group.

To conclude, in order to take full advantage of these preliminary results, further lengthy data analysis is needed. The authors of this paper anticipate that the final results of this post-processing stage will provide a rich basis for another publication.

Acknowledgment

The research leading to these results has received EU funding under grant agreement n° 600796 (MOBOT project, FP7-ICT).

REFERENCES

[1] M. Hirvensalo, T. Rantanen, and M. Heikkinen, "Mobility difficulties and physical activity as predictors of mortality and loss of independence in the community-living older population," Journal of the American Geriatric Society, vol. 48, pp. 493-498, 2005.

[2] O.J. Chuy, Y. Hirata, Z. Wang, and K. Kosuge, "Approach in assisting a sit-to-stand movement using robotic walking support system," IEE/RSJ Int. Conf. on Intelligent Robots and Systems, China, 2006.

[3] D. Chugo, T. Asawa, T. Kitamura, S. Jia, and K. Takase, "A rehabilitation walker with standing and walking assistance," IEE/RSJ Int'l Conf. on Intelligent Robots and Systems, Nice, France, 2008.

[4] K. Wakita, J. Huang, P. Di, K. Sekiyama, and T. Fukuda, "Human walking-intention-based motion control of an omnidirectional-type cane robot," In IEEE/ASME Transactions On Mechatronics, vol. 18, no. 1, 2013.

[5] Y. Hirata, S. Komatsuda, and K. Kosuge, "Fall prevention control of passive intelligent walker based on human model," IEE/RSJ Int'l Conference on Intelligent Robots and Systems. Nice, France, 2008.

[6] S-E. Fotinea, et al., "The MOBOT Human-Robot Communication Model", in Proc. of 6th IEEE CogInfoCom 2015 Conference, 19-21 October 2015, Győr, Hungary, pp. 201 – 206.

[7] X. Papageorgiou, et al., "Advances in intelligent mobility assistance robot-integrating multimodal sensory processing," Universal Access in Human-Computer Interaction. Aging and Assistive Environments, C. Stephanidis and M. Antona Eds.: UAHCI/HCII 2014, Part III, Lecture Notes in Computer Science, Switzerland: Springer International Publishing, pp. 692-703, 2014.

[8] S-E. Fotinea, et al., "Data acquisition towards defining a multimodal interaction model for human-assistive robot communication," C. Stephanidis and M. Antona Eds.: UAHCI/HCII 2014, Part III, Lecture Notes in Computer Science, Switzerland: Springer International Publishing, pp. 615-626, 2014.

[9] E. Trulls, I. Kokkinos, A. Sanfeliu, and F. Moreno, "Superpixel-grounded deformable part models," IEEE Conf. on Computer Vision and Pattern Recognition, CVPR, 2014.

[10] H. Boussaid, I. Kokkinos, and N. Paragios, "Discriminative learning of deformable contour models," Int. Symposium on Biomedical Imaging (ISBI), 2014.

[11] E. Trulls, I. Kokkinos, A. Sanfeliu, and F. Moreno, "Dense segmentation-aware descriptors," IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2013.

[12] V. Pitsikalis, A. Katsamanis, S. Theodorakis, and P. Maragos, "Multimodal gesture recognition via multiple hypotheses rescoring," Journal of Machine Learning Research, pp. 255-284, 2015.

[13] N. Rodomagoulakis, et al., "Multimodal human action recognition in assistive human-robot interaction," ICASSP China, 2016 [Proceedings International Conf. on Acoustics, China, 2016].

[14] E. Efthimiou, et al. "The MOBOT platform-showcasing multimodality in human-assistive robot interaction," in International Conference on Universal Access in Human-Computer Interaction. Springer International Publishing, 2016.

[15] S. Escalera, et al., "Multi-modal Gesture Recognition Challenge 2013: Dataset and Results," ICMI 2013.

[16] Z.I. Skordilis, et al., "Multichannel speech enhancement using MEMS microphones," IEEE Australia, 2015 [Int'l Conf. on Acoustics, Speech, and Signal Processing, Australia, April 2015.

[17] J. Corredor, J. Sofrony, and A. Peer, "Deciding on optimal assistance policies in haptic shared control tasks," IEEE 2014 [Int'l Conference on Robotics and Automation, pp. 2679-2684, 2014].

[18] K.-L. Ho Hoang, D. Corradi, and K. Mombaur, "Identification and classification of geriatric gait patterns - multi-contact capturability," French-German-Japanese Conference in Humanoid and Legged Robots, May 2014, Germany, 2014.

[19] X.S. Papageorgiou, G. Chalvatzaki, C.S. Tzafestas, and P. Maragos, "Hidden Markov modeling of human normal gait using laser range finder for a mobility assistance robot," IEEE China, June 2014 [Int'l Conference on Robotics and Automation, pp. 482-487, 2014].

[20] MOBOT Deliverable D5.3 - Report on performance metrics and first evaluation study: http://www.mobot-project.eu/userfiles/downloads/Deliverables/MOBOT_WP5_D5.3.pdf.

[21] MOBOT Deliverable D5.1: Preliminary report on use cases and user needs: http://www.mobot-project.eu/userfiles/downloads/Deliverables/ MOBOT_WP5_D5.1_v1.5.pdf.

[22] MOBOT Deliverable D5.2 : Report on use cases, performance metrics and user study preparations: http://www.mobot-project.eu/userfiles/downloads/Deliverables/MOBOT_WP5_D5.2_v1.5.pdf.

[23] MOBOT Deliverable D5.4: Report on performance metrics and final evaluation study: http://www.mobot-project.eu/userfiles/downloads/ Deliverables/ MOBOT_WP5_D5.4.pdf.

[24] R. Kachouie, S. Sedighadeli, R. Khosla, and M.T. Chu, "Socially assistive robots in elderly care: a mixed-method systematic literature review," International Journal of Human-Computer Interaction, vol. 30.5, pp. 369-393, 2014.

[25] D.O. Johnson, et al. "Socially assistive robots: a comprehensive approach to extending independent living," International Journal of Social Robotics, vol. 6.2, pp. 195-211, 2014.

[26] C. Werner, P. Ullrich, M. Geravand, A. Peer, and K. Hauer, "Evaluation studies of robotic rollators by the user perspective: a systematic review," Gerontology 2016.

[27] L. Schwickert, J. Klenk, A. Stahler, C. Becker, U. Lindemann, "Robotic assisted rehabilitation of proximal humerus fractures in virtual environments:a pilot study" Z Gerontol Geriatr,vol.44,pp.387-392, 2011.

[28] J. Cohen, Statistical Power Analysis for the Behavioral Sciences, Hilsdale, Lawrence Erlbaum Associates, 1988.

[29] F. Rudzicz, R. Wang, M. Begum and A. Mihailidis, "Speech interaction with personal assistive robots supporting aging at home for individuals with Alzheimer's disease," ACM Transactions on Accessible Computing (TACCESS), 7.2: 6, 2015.

[30] A. Kötteritzsch, and B. Weyers, "Assistive technologies for older adults in urban areas: a literature review," Cognitive Computation, vol. 8.2, pp. 299-317, 2016.

[31] I. Rodomagoulakis, et al., "Multimodal human action recognition in assistive human-robot interaction," IEEE Int'l Conference on Acoustics, Speech and Signal Processing (ICASSP), 2016.

[32] C. Feichtenhofer, A. Pinz, and A. Zisserman, "Convolutional two-stream network fusion for video action recognition," arXiv preprint arXiv:1604.06573, 2016.

[33] Y. Koumpouros, "A systematic review on existing measures for the subjective assessment of rehabilitation and assistive robot devices," Journal of Healthcare Engineering, vol. 2016, Article ID 1048964, 10 pag Hindawi Publishing Corporation, 2016.

[34] Y. Koumpouros, E. Papageorgiou, and A. Karavasili, "Development of a new psychometric scale (PYTHEIA) to assess the satisfaction of users with any assistive technology," in Advances in Human Factors and Ergonomics in Healthcare, Advances in Intelligent Systems and Computing 48,2V.G. Duffy and N. Lightner, Eds., Switzerland: Springer International Publishing, 2017, pp. 343-353.

[35] Y. Koumpouros, E. Papageorgiou, A. Karavasili, and F. Koureta, "PYTHEIA: A scale for assessing rehabilitation and assistive robotics," ICPMRSM 2016: 18th Int'l Conference on Physical Medicine, Rehabilitation and Sports Medicine, November 2016, South Africa.

[36] Y. Koumpouros, A. Karavasili, E. Papageorgiou, and P. Siavelis, "Validation of the Greek version of the device subscale of the Quebec user evaluation of satisfaction with assistive technology 2.0 (QUEST 2.0)," Assistive Technology: The official Journal of RESNA, 2016.

[37] Y. Koumpouros, E. Papageorgiou, A. Karavasili, and D. Alexopoulou, "Translation and validation of the assistive technology device predisposition assessment in Greek in order to assess satisfaction with use of the selected assistive device," Disability and Rehabilitation: Assistive Technology, 2016.