Autonomous Function of Wheelchair-Mounted Robotic Manipulators to Perform Daily Activities

Cheng-Shiu Chung MS [1,2],
[1]Human Engineering Research Laboratories
Department of Veterans Affairs
Pittsburgh, PA [2] Department of Rehabilitation Science and Technology
University of Pittsburgh

Joshua.Chung.CS@gmail.com

Hongwu Wang PhD [1,2]
[1]Human Engineering Research Laboratories
Department of Veterans Affairs
Pittsburgh, PA
[2] Department of Rehabilitation Science and Technology
University of Pittsburgh
How11@pitt.edu

Rory A. Cooper PhD [1,2]
[1]Human Engineering Research Laboratories
Department of Veterans Affairs
Pittsburgh, PA
[2] Department of Rehabilitation Science and Technology
University of Pittsburgh

rcooper@pitt.edu

***Abstract*—Autonomous functions for wheelchair-mounted robotic manipulators (WMRMs) allow a user to focus more on the outcome from the task – for example, eating or drinking, instead of moving robot joints through user interfaces. In this paper, we introduce a novel personal assistive robotic system based on a position-based visual servoing (PBVS) approach. The system was evaluated with a complete drinking task, which included recognizing the location of the drink, picking up the drink from a start location, conveying the drink to the proximity of the user's mouth without spilling, and placing the drink back on the table. For a drink located in front of the wheelchair, the success rate was nearly 100%. Overall, the total time of completing drinking task is within 40 seconds.**

Keywords—object recognition; wheelchair; robotics; manipulation; path planning

I. Introduction

Control of the user interface may be the major barrier to using a wheelchair-mounted robotic manipulator (WMRM) for people with upper extremity impairments when performing activities of daily living (ADL). The user interface is supposed to allow the user to focus more on the tasks to be performed instead of how to control the robot [1]. For example, the focus of eating and drinking tasks should be on the pleasure of food and drink instead of how to convey the food or drink to one's mouth.

There are two ways to overcome this barrier: Image-Based Visual Servoing (IBVS) and Position-Based Visual Servoing (PBVS) [2]. In the IBVS, the gripper is guided toward an object at the region of interest through feedback from a camera mounted on the robot gripper or wrist. The IBVS has been integrated into WMRMs for pick-and-place task by University of South Florida [3] and University of Massachusetts Lowell [4]. The object location is continuously recognized by the image in order to adjust robotic movement until the gripper picks up the object. The PBVS uses images to localize the object in 3D space, and plans a trajectory toward the object with knowledge of the environment [2]. The camera can be mounted on the wheelchair or robot shoulder to provide a better perspective of the object and its surroundings. The major difference between these two approaches is the handling of occlusion and starting location of movement. The PBVS has the advantage of finding a path and grasping plan even when the object is occluded from the starting location or folding position [5]. Currently, there is no PBVS integrated into WMRMs.

In this paper, we introduce a novel personal assistive robotic system which is based on the mechanical design of the Personal Mobility and Manipulation Appliance (PerMMA) developed by the University of Pittsburgh [6] and the software architecture of the Home Exploring Robotic Butler (HERB) developed by Carnegie Mellon University [5]. This system is based on the PBVS approach using Multiple Object Pose Estimation and Detection (MOPED) to identify the distance and pose of objects and Constraint Bidirectional Rapid Random Tree (CBiRRT) as the path planning strategy. The system can be evaluated by a complete drinking task which includes picking up the drink from a starting location, conveying the drink to the proximity of the user's mouth without spilling, and placing the drink back on the table.

II. Related Works

The mechanical system was designed similarly to the PerMMA robot. It is composed of two moveable WMRMs

Research supported by the National Science Foundation, the US Department of Veterans Affairs and the National Institute for Disability and Rehabilitation Research.

Hongwu Wang and Cheng-Shiu Chung are with the Human Engineering Research Laboratories, VA Pittsburgh Healthcare System and University of Pittsburgh, Pittsburgh, PA 15206 USA.

Rory A Cooper. Author is with the Human Engineering Research Laboratories, VA Pittsburgh Healthcare System and University of Pittsburgh, Pittsburgh, PA 15206 USA (phone: 412-822-3700; fax: 412- 822-3699; e-mail: rcooper@pitt.edu).

mounted on a track located around the wheelchair seat so that the manipulators can slide to the back of the chair while driving through a narrow hallway or door. Three modes of the user interface were developed: local user, remote user, and cooperative control mode. Local user mode allows the wheelchair user full WMRM control on PerMMA by using a touchpad or via speech recognition. Remote user mode shifts the authority of WMRM control to a remote operator, who could be a caregiver or family member of the wheelchair user. In this way, caregivers can remotely complete ADL through the visual feedback from the cameras on the robot shoulders. Cooperative mode takes the advantages of better perception form the local wheelchair user and the better dexterity from the remote operator. A user study was conducted in performing ADL with PerMMA using keyboard, keypad, and remote operation and voice control and touch screen interfaces are under clinical trial. A previous focus group study reveals that users prefer to control PerMMA by themselves [6]. However, with the complexity of the system, it is difficult for all users to use the current interfaces to perform ADL. The long-term goal of current work is to integrate autonomous functions with other interfaces to provide smart assistance. The current work is the first approach to develop autonomous functions for PerMMA.

The WMRM used in this novel system is an iARM, manufactured by Exact Dynamics (Didam, the Netherlands). The iARM is a six degree-of-freedom (DOF) robotic arm with two-fingered hand, and is an upgraded version of the Manus ARM. The iARM is lighter than Manus ARM mainly due to removal of gravity compensation springs. It can be controlled by keypad, joystick, or single-button switches, and can be mounted on a powered wheelchair with a camera mounted on its shoulder to detect objects [7], [8].

The object recognition algorithm used in this system is MOPED, which is reliable and robust in complex environments with low latency [9]. The pose and distance of the object can be estimated by a single image. The image is first processed by extracting features with Scale-Invariant Feature Transform (SIFT). The extracted features are compared with the stored SIFT features using an offline learning procedure. The matched features are clustered by Iterative Clustering Estimation, which iteratively uses Random Sample Consensus or Levenberg-Marquardt to estimate the object pose hypotheses. These pose estimations are clustered with an implemented object hypothesis scoring function based on M-estimator theory to eliminate the outliers. By taking the advantage of parallel computation of GPU/CPU hybrid architecture, low latency can be achieved [9].

Following the estimation of the pose and location of the objects, the path planning for manipulation in this system is Constrained Bidirectional Rapid Random Tree (CBiRRT). This planning algorithm is composed of three components: constraint representation, constraint-satisfaction strategies, and a general planning algorithm. The constraint can be represented using Task Space Regions (TSRs) representation. TSRs are representation of pose constraints that can be described based on the tasks. Moreover, TSRs are also capable of linking together for complex tasks or end-effector poses. For example, two TSRs are used while bringing the drink to the user. One TSR is to define the acceptable space that the drink will be conveyed to. Another TSR is to keep the drink upright all the time during the movement [10].

A study monitored the ADLs of an able-bodied participant for five days and identified 3964 activities based on the International Classification of Functioning, Disability and Health (ICF) [11]. Among these activities, the most frequent task for self care (d5 – ICF code) is drinking. The drinking task includes several of the most frequent mobility tasks for carrying, moving, and handling objects (d430 – d449) such as lifting (d4300), putting down objects (d4305), manipulating (d4402), and carrying in the hands (d4301). Therefore, in this paper, we have evaluated our system using a drinking task, which is one of the most frequently used self care daily tasks that requires complex manipulation skills.

III. System Architecture

A. Hardware

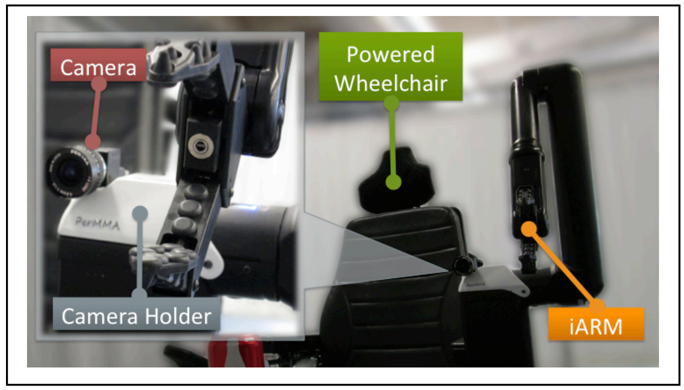

Fig. 1. Picture of the hardware of the system

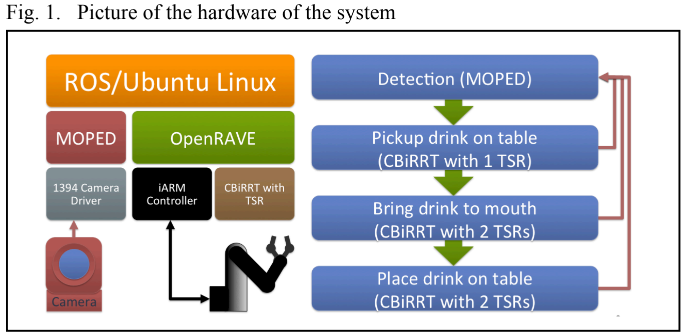

Fig. 2. Software architecture of the system (left) and state flow of the testing procedure (right)

Similar to PerMMA's mechanical design, the iARM is mounted on the side of a powered wheelchair. An IEEE-1394 fire-wire camera (Flea 2), manufactured by Point Grey (Richmond, British Columbia, Canada), is mounted with an in-house manufactured holder. A Pentax TV lens with wide field of view (4.8mm 1:1.8) is attached to the camera. The camera is connected to a Lenovo laptop (CPU: 8-core i7-2960XM, RAM: 16GB, GPU: Quadro 1000M, running Ubuntu Linux 10.04) for the processing of object recognition and path planning. The iARM is connected to a CAN Bus/serial module and linked to the laptop through an USB/serial conversion cable. Figure 1 shows the picture of the system.

B. Software Architecture

Inherited from HERB's software design [5], the software can be described as having the following structure: sensing, planning, performing. The system first recognizes the object's pose and location in the environment. The planning algorithm then searches for an optimized trajectory to pick up the object under the environmental geometries and WMRM kinematics with constraints. The trajectory is then performed on the robot to physically retrieve the object.

The software architecture of this system was designed with two major features: expandable computational power and minimal human input for sensing and planning algorithms. The communication infrastructure and process of the sensing and planning algorithms are managed by the Robot Operating System (ROS) package, which also provides the capability of transferring computational processing between computers. The remote operator can also control the robot under this software structure. The touch screen control user interface can also be integrated into a mobile phone or tablet [5].

The MOPED is capable of detecting multiple objects and estimating their 3D poses and locations using a 640×480 grayscale image. The detected objects are automatically placed into the OpenRAVE simulation environment. The OpenRAVE environment conducts path planning (CBiRRT), simulates robotic motions, and generates the trajectory. The trajectory is then sampled to several waypoints that contain the joint angles and velocities. The waypoints are sent to the iARM joint position control function to move through them. The driver for the iARM control was also developed to publish joint angular messages and provide robotic movement services through ROS.

IV. TESTING PROCEDURE

A drinking task was used for the preliminary evaluation of the autonomous function of the system for two reasons. The first reason is that, as previously described, this is the most frequent daily self care task according to [11]. The second reason is the complexity of the task. People without upper extremity impairments perform this task in seconds without thinking about the motion of their arm and hand. However, manually controlling the WMRM using the touchpad or speech recognition requires people with upper extremity impairments to combine three-dimensional vision in locating the object and two-dimensional movements across the tabletop, plus a "grasp" command to either grasp and lift or tilt the robotic hand. During the movement, the arm or gripper may occlude the target object from being grasped. These barriers make it difficult and time consuming to get a sip of drink. As a result, making an autonomous function for a drinking task not only saves time but also increases the quality of self-care.

A. Subtasks

The drinking task is simplified into four subtasks: detection of the drink, planning and pickup of the drink on the table, bringing the drink to the proximity of the user, and placing it back onto the table. We used a soda can as the drink for this task since it is a common drink, but the algorithms also work for other kinds of drinks. Python scripts were developed to control the flow of states in the drinking task and manage error recovery strategies. The state flow is shown in Figure 2. Any failure during the movement subtasks was recorded and treated as a failed trial. In addition, for the safety of the occupant, a trial with any collision with the wheelchair user will be rated as a failed trial.

Two starting locations as shown in Figure 4, easy and difficult, were used to evaluate the capability of the system in handling difficult tasks. The easy start location had the gripper above the table with no occlusion between the gripper and the soda can. In the difficult start location, the gripper started under the table. In the difficult configuration, the table is a long and large occlusion between the gripper and drink, and there is only a 3cm gap between the WMRM's elbow joint and the table.

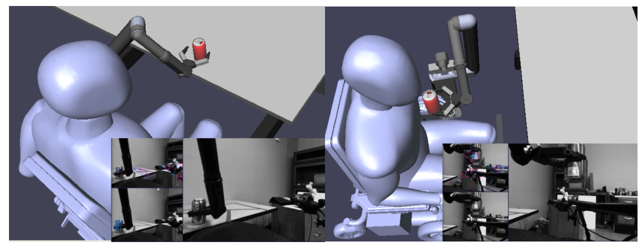

Fig. 3. The subtask of pickup (left) and drink (right)

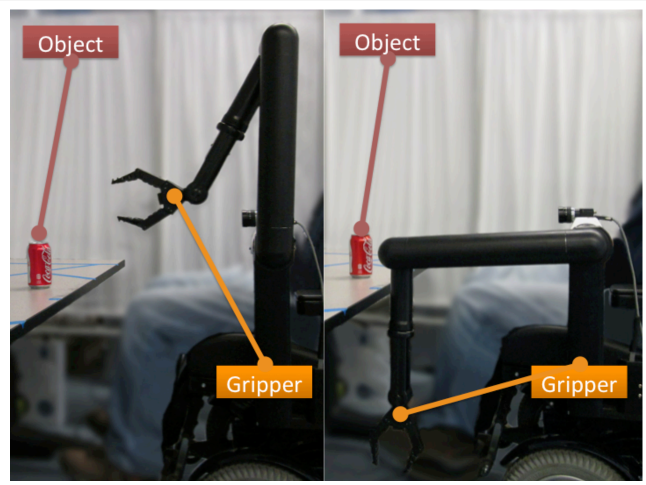

Fig. 4. Two start locations of the pickup subtask. The higher location (left) and the lower location (right)

B. Parameters for CBiRRT

The TSR defines the constraints that limit the CBiRRT path planning from searching unwanted trajectories or unwanted end-effector poses. The Bound B^w in the TSR is defined as (1).

$$B^w = \begin{bmatrix} x_{min}, x_{max} & y_{min}, y_{max} & z_{min}, z_{max} \\ yaw_{min}, yaw_{max} & pitch_{min}, pitch_{max} & roll_{min}, roll_{max} \end{bmatrix} \quad (1)$$

The min and max indicate the lower and upper boundary of the constraint. For example, the B^w = [0,0; 0,0; 0,0; 0,0; 0,0; -π/2,π/2] indicates that there is no freedom in the xyz direction as well as the yaw and pitch angles but the roll angle allows rotation from -90 to 90 degrees. Another example of B^w = [-100,100; -100,100; -100,100; 0,0; 0,0; -π,π] represents that the xyz directions and roll angle allow movement but not the yaw and pitch angle.

For the subtask of picking up the soda can, we only constrained the end-effector. However, for the other subtasks of conveying the drink, there was one more constraint applied for preventing the drink from spilling. The parameters applied are listed in Table 1. We set the time limit for searching end-effector solutions to 5 seconds and the time for searching the entire trajectories to 30 seconds for each subtask. The iterations number for smoothing trajectory is 150.

TABLE I. TSR PARAMETERS OF THE SUBTASKS

Subtask	B^w	Bw Type
Pickup the drink	[0,0;0,0;0,0;0,0;0,0; -π/2,π/2] if the drink is at the left hand side of the iARM gripper	Goal pose
	[0,0;0,0;0,0;0,0;0,0; π/2,3π/2] if the drink is at the right hand side of the iARM gripper	Goal pose
Bring the drink to the user	[-0.35, -0.4, 1.05]	User's mouth
	[0,0;0,0;0,0;0,0;0,0; -π/2,π/6]	Goal pose
	[-100,100;-100,100;-100,100; 0,0;0,0; -π ,π]	Constrain
Place the drink on the table	[0,0;0,0;0,0.1;0,0;0,0;-π/2,π/2]	Goal pose
	[-100,100;-100,100;-100,100; 0,0;0,0; -π ,π]	Constrain

C. Evaluation Measures

The system was tested on three levels: the detection level, the planning level, and the whole system level. The detection level only evaluates the MOPED in detecting different orientations of the soda can. The soda can faces the camera with different rotation angles and distances. The planning level (shown in Figure 5) evaluates the ability and success rate of the CBiRRT with OpenRAVE simulation in searching trajectories for each subtask. In this test, the soda can was randomly placed in front of the WMRM either inside or outside its working space. The whole system level test evaluates the system, including moving the iARM to physically picking up the drink and bringing it to the user, to see how the planning strategies work in the real world. As shown in Figure 4, the soda can was randomly put on the table inside the working space and recognition area (inside the blue tape).

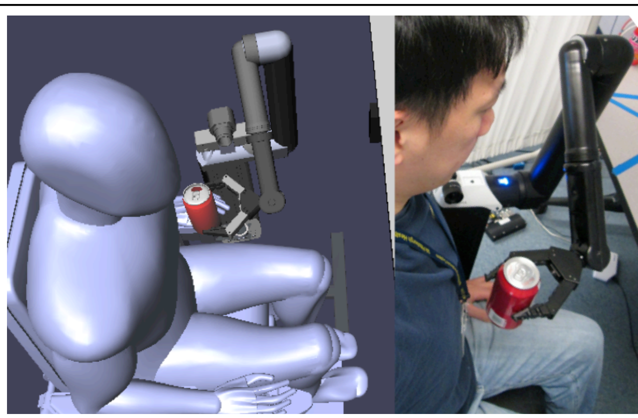

Fig. 5. The planning level (right) and system level (left)

The success rate and completion time are the major outcome measures for each subtask. The time of detection is determined by the start of the detection state to successfully finding the drink. For the path planning level, the success rate of planning algorithms and the time needed for planning is reported, including the robot simulation. However, we added one more condition that if the time of planning and robot simulation exceeds 60 seconds, i.e. slower than human performance [6], we rate this trajectory as failed. The location failures to find trajectories were also recorded. The average speed was defined by equation (2).

$$Speed = \frac{\sqrt{(x_T-x_0)^2+(y_T-y_0)^2+(z_T-z_0)^2}}{T} \quad (2)$$

where T is the time from the start of path planning to the end of robot movement. x_T, y_T, z_T are the position at the end of the trajectory and x_0, y_0, z_0 are the start position. Table II shows the outcome measures for each subtask.

TABLE II. MEASUREMENTS OF SUBTASKS

Subtask	Measurements
Detection	Successful rate Time completion
Pickup the drink	Successful rate Time completion/Speed Fail reasons
Bring the drink to the user	Successful rate Time completion/Speed Fail reasons
Place the drink on the table	Successful rate Time completion/Speed Fail reasons

V. RESULTS

The results of completion time, moving speed, and the success rate are shown in Table III.

TABLE III. TEST RESULTS

Subtask		Outcome measure		
		Complet. time (second)[a]	*Speed (mm/s)*	*#Fail/#Total Success Rate*
Planning level (start above the table)	Pickup	3.62±0.80 (1.32~7.70)	134.6±41.4 (55.4~358.0)	5/1365 99.6%
	Drink	2.51±0.99 (0.86~5.95)	339.2±121.9 (141.1~777.8)	47/1127 96.17%
	Place	1.83±0.47 (0.80~4.67)	436.3±95.2 (168.5~847.4)	0/1170 100%
Planning level (start under the table)	Pickup	9.50±9.74 (1.93~58.53)	79.3±39.1 (6.4~210.4)	18/558 96.8%
	Drink	2.6±0.9 (1.0~5.4)	322.6±115.0 (139.7~730.6)	37/472 92.2%
	Place	1.9±0.5 (0.9~4.6)	425.9±94.0 (158.1~782.1)	0/505 100%
System level	Detect	0.45±0.12 (0.20~0.68)	N/A	100% (0°) 92% (45°)
	Pickup	12.1±2.6 (7.6~18.0)	47.5±5.8 (26.8~50.7)	18/62 70.1%
	Drink	9.58±1.85 (6.6~15.4)	74.6±12.3 (50.3~99.3)	7/38 81.6%
	Place	10.6±2.7 (5.4~17.0)	75.8±17.4 (48.1~125.8)	0/30 100%

[a] Completion time is presented as average ± standard deviation (minimum ~ maximum)

A. Detection Performance

The average time for MOPED detection was 0.45 second for a single soda can and 1.75 seconds for multiple objects shown in Figure 6. However, the faces with less SIFT features were harder to recognize (Shown in Figure 6). Detection success rates were 100% at 0 degrees, 92% at 45 degrees, and unable to identify at 90 degrees. Moreover, the MOPED distance estimation was 1 inch shorter when the soda can was more than 28 inches away.

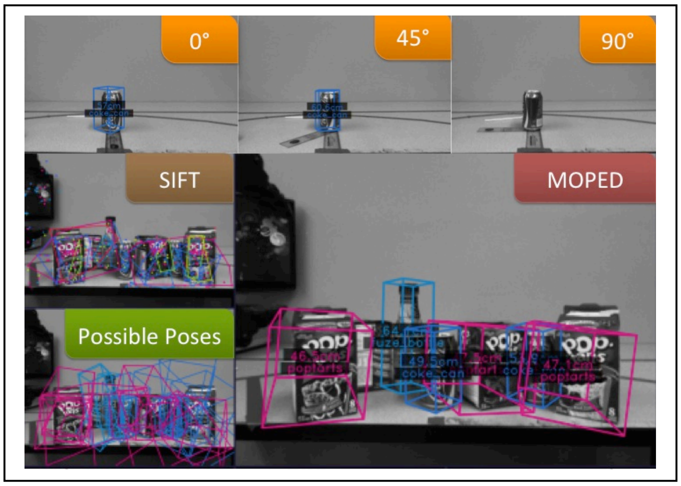

Fig. 6. Different faces and poses of soda recognized by MOPED

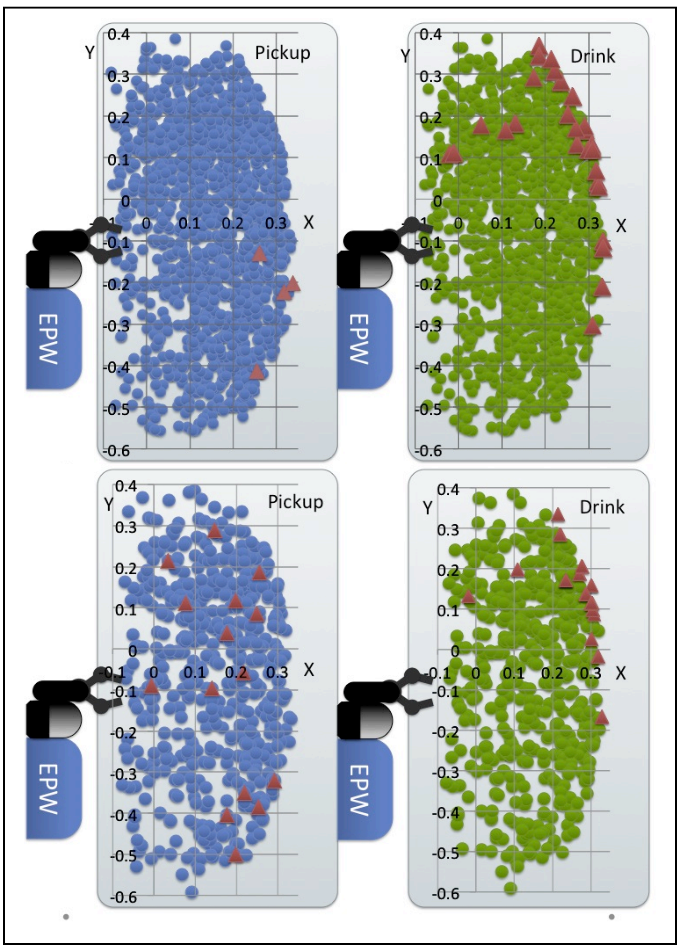

Fig. 7. Success and fail location on the table (upper left: Pickup subtask from higher start location, upper right: Drinking subtask from higher start location, lower left: Pickup subtask from lower start location, lower right: Drinking subtask from lower start location)

B. Movement Performance

Overall, the path planning simulations show a very high success rate (> 92%). The pickup subtask is relatively slower than the drink and place subtasks. Planning from the easy location was faster than from the difficult position. Figure 7 plots the locations on the table that have been tested for pick-up and drinking subtasks. The red triangles indicate the location from which the iARM was unable to complete the subtask. The iARM and electrical power wheelchair (EPW) is drawn on the side of the table. In the pick-up subtask, starts from under the table show more random failures than starts from above the table. More failures were located on the left side of the iARM. The failures on the right side of the iARM were close to the limit of the workspace. In the drinking subtask, most of the failures were found at the edge of the iARM workspace. There were no failures found on the right side of the iARM.

In the system test, the overall success rate was 70.1% for the pick-up subtask and 81.6% for the drinking subtask. The speed of the pick-up subtask was slower than the simulation. The completion time of the WMRM in the subtasks was longer than the simulation (Table III). The entire drinking task was completed within 40 seconds. The successful and failed locations on the table of four subtasks are plotted in Figure 8.

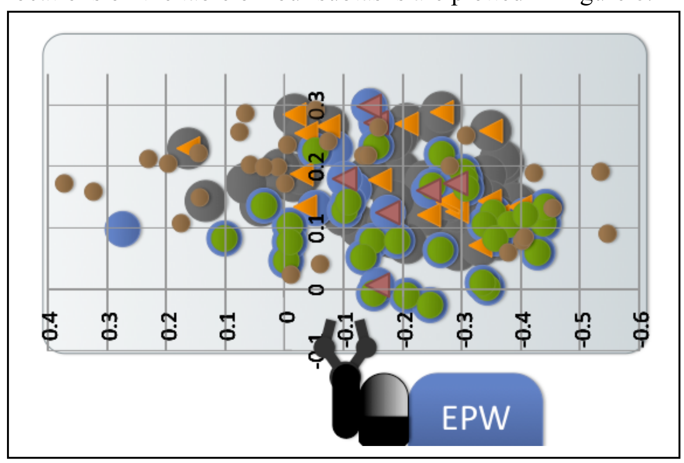

Fig. 8. The successful and failed locations on the table of four subtasks in the system level test. Gray: success detection; Blue: successful pickup; Orange: failed pickup; Green: successful drink; Red: failed drink; Brown: place locations.

VI. Discussion

A. Detection

The MOPED relies on the SIFT features for establishing object pose hypotheses. The object with fewer SIFT features has less chance of being recognized. The other limitation was the calibration of the camera. Although the camera was calibrated before tests to eliminate distortion and skew factors form the lens, the error in the camera internal parameters may be amplified if the object was away from the camera. Therefore, at the edge of the workspace, the estimated distance error is about 1 inch.

B. Movement

The planning simulations demonstrated a very high success rate at the easy and difficult start locations. Most of the planning failures occurred at the edge of the workspace. This was probably because of the singularity point in the kinematic model when the robot is fully extended. For the subtask of picking up and bringing the drink to the user from the difficult location, the robot failed more often on its left hand side (0.2m away from robot base). This is similar to a human's arm, in that it is harder for a human to bring an object far away from the body. The robot had no problems in picking up and bringing to the user when the soda can was in front of the wheelchair and about 0.3m from the edge of the table.

In the system level test, the major cause that the pick-up failed was when the MOPED positioning error was located at the far end away from the camera. The failures were likely caused by the aggressive trajectories. These kinds of trajectories include some motions with either 0.5" tolerance to the objects or arm extended more necessary. The drinking subtasks usually failed with unsafe grasping that dropped the object during motion. These failed moves can be improved with better trajectory strategies. Although the speed was slower than the simulation, it can be increased by re-sampling the trajectory to fewer waypoints so that the WMRM has fewer stops during the movement.

VII. CONCLUSION

A novel WMRM system with autonomous functions was developed. This system was evaluated with a drinking task that included carrying and handling the drink. The drinking task was divided into four subtasks: detection, picking up the drink, bringing the drink to the user, and placing the drink on the table. Success rates and the average task completion time for each subtask were computed. The entire drinking task was completed within 40 seconds.

Future work includes error correction implementation of object recognition to reduce the distance error, path planning strategies to reduce failures in performing on the WMRM, and optimization of the trajectory waypoints to maximize the speed of the WMRM. Other work will include a user interface for picking up object autonomously and combination with local user interface for cooperative mode control.

ACKNOWLEDGMENT

The authors would like to thank Mehmet Dogar, Alvaro Collet, Kyle Strabala, Anca Dragan, Christopher Dellin, Mike Vande Weghe, Dr. Srinivasa, and Annmarie Kelleher for their valuable help on this work. This material is based upon work supported in part by the National Science Foundation Grant#EEC-540865and the Department of Veterans Affairs, Rehabilitation Research and Development Service (B6789C). The contents of this paper do not represent the views of the Department of Veterans Affairs or the United States Government.

REFERENCES

[1] W. S. Harwin, T. Rahman, and R. A. Foulds, "A Review of Design Issues in Rehabilitation Robotics with Reference to North American Research," *Rehabilitation Engineering, IEEE Transactions*, vol. 3, no. 1, pp. 3–13, 1995.

[2] K. Jablokow, "Visual control of robots: High-performance visual servoing," *Control Engineering Practice*, vol. 5, no. 9, p. 1337, Sep. 1997.

[3] W. Pence and F. Farelo, "Visual servoing control of a 9-DoF WMRA to perform ADL tasks," in *2012 IEEE International Conference on Robotics and Automation*, 2012, pp. 916–922.

[4] K. M. Tsui, D.-J. Kim, A. Behal, D. Kontak, and H. A. Yanco, "I Want That ": Human-in-the-Loop Control of a Wheelchair-Mounted Robotic Arm," *Applied Bionics and Biomechanics*, vol. 8, no. 1, pp. 127–147, 2011.

[5] S. S. Srinivasa, D. Ferguson, C. J. Helfrich, D. Berenson, A. Collet, R. Diankov, G. Gallagher, G. Hollinger, J. Kuffner, and M. Vande Weghe, "HERB: a home exploring robotic butler," *Autonomous Robots*, vol. 28, no. 1, pp. 5–20, Nov. 2009.

[6] H. Wang, G. G. Grindle, J. Candiotti, C. Chung, M. Shino, E. Houston, and R. a Cooper, "The Personal Mobility and Manipulation Appliance (PerMMA): A robotic wheelchair with advanced mobility and manipulation.," in *34th Annual International Conference of the IEEE EMBS*, 2012, vol. 2012, pp. 3324–7.

[7] S. W. Brose, D. J. Weber, B. A. Salatin, G. G. Grindle, H. Wang, J. J. Vazquez, and R. A. Cooper, "The role of assistive robotics in the lives of persons with disability.," *American journal of physical medicine & rehabilitation / Association of Academic Physiatrists*, vol. 89, no. 6, pp. 509–21, Jun. 2010.

[8] J. W. Capille, S. Carey, R. M. Alqasemi, and R. Dubey, "Kinematic Evaluation of Commercial Wheelchair-Mounted Robotic Arms," in *2011 IEEE International Conference on Systems Man and Cybernetics*, 2011, pp. 711–716.

[9] A. Collet, M. Martinez, and S. S. Srinivasa, "The MOPED framework: Object recognition and pose estimation for manipulation," *The International Journal of Robotics Research*, vol. 30, no. 10, pp. 1284–1306, Apr. 2011.

[10] D. Berenson, S. Srinivasa, and J. Kuffner, "Task Space Regions: A framework for pose-constrained manipulation planning," *The International Journal of Robotics Research*, vol. 30, no. 12, pp. 1435–1460, Mar. 2011.

[11] Y. Matsumoto, Y. Nishida, Y. Motomura, and Y. Okawa, "A concept of needs-oriented design and evaluation of assistive robots based on ICF.," in *IEEE International Conference on Rehabilitation Robotics*, 2011, vol. 2011, p. 5975437.