A Framework for Door Localization and Door Opening Using a Computer Controlled Wheelchair for People Living with Mobility Impairments

Alexander Andreopoulos, John K. Tsotsos
Department of Computer Science and Engineering
Centre for Vision Research
York University
Toronto, Ontario, Canada
Email: {alekos, tsotsos}@cse.yorku.ca

***Abstract*— We present a framework for localizing and opening doors that fall within our current field of view, using Playbot, a computer controlled wheelchair that is equipped with vision sensors and a 6+2 degrees of freedom robotic arm. Contributions of the paper include: (i) Reliable detection of a specular door handle from close distances, where intense specularities tend to make many object detection algorithms unreliable (ii) Reliable stereo vision depth extraction of a specular door handle and subsequent opening of the door using a 6+2 degrees of freedom robotic arm. We present results demonstrating the validity of the approach and we discuss promising directions for future research.**

I. Introduction

With a rapidly aging population in the developed world, service robots that can help people with mobility impairments lead a more independent and productive life, are becoming ever more important [5]. Playbot is a long-term large-scale research project whose goal is to provide a computer controlled wheelchair which may enable children and adults with mobility impairments become more independent [14]. Tasks that we want Playbot to perform include the ability to visually search the environment (Active Visual Search [12, 15, 19]), to recognize objects and events and to work in natural dynamic and unpredictable environments. The research of the project is focused on Playbot's vision, as this is the major bottleneck to the development of intelligent robots [14]. Therefore, vision is the primary sensor on Playbot and we do not use other sensors such as laser range finders or sonars.

A number of researchers have previously dealt with the door opening problem using a mobile manipulator. In [9] a control strategy for door opening with the use of a mobile 3-fingered manipulator is proposed. Active sensing algorithms are proposed to overcome uncertainties in a real environment. Rather than using a wrist force/torque sensor for force and position control, the contact force data of a multi-fingered robot hand is used. In [8] the authors implement a door opening controller using a hybrid dynamic system model that is more abstract than traditional continuous control techniques and uses relaxation of force control. The results demonstrate that higher level models for controlling the manipulator can lead to significantly lower error during the door handle grasping task. In [7] the authors use the path of least resistance to control an arm and open a door. In [6] the authors model the door opening task using a sequence of planned motion primitives – called "action primitives" by the authors – where each action primitive is designed with an error adjustment mechanism to help deal with positioning errors of the mobile base. The authors assume that the position and radius of the door is known beforehand. In [2] Brooks *et al.* present a robot equipped with a dexterous arm that is capable of finding a door, pushing it open and going through it. However, the robot does not deal with the problem of reaching and turning the door handle.

In this paper we present a framework used by Playbot for localizing and opening doors. The contributions include: (i) Reliable detection of a specular door handle from close distances, where wild intensity fluctuations due to specularities tend to make many object detection algorithms unreliable (ii) Reliable depth extraction of a specular door handle and subsequent opening of the door using a $6+2$ degrees of freedom robotic arm and stereo vision.

In Section II we overview the Playbot project within the context of which the door lozalization and opening behaviors were built. In Section III we describe our approach for localizing door handle instances and subsequently opening the door using a robotic arm. In Section IV we present experimental results of the localization performance and the door opening performance demonstrating the reliability of our approach. We conclude the paper by discussing promising topics for future research within the context of the Playbot project.

II. The Playbot Project

Current assistive technology for the physically disabled rely on the user's visual system as part of a closed-loop control system. For example, in one class of robotic aids, specialized sensors are developed for a finger or eyebrow. To grasp an object, the user visually guides a robot manipulator through a series of micro-activations to the target. Each micro-activation moves a particular joint of a robot arm by a small distance.

Fig. 1. The Playbot wheelchair which we used in our experiments.

This can be tedious, especially for children, as the user tires easily and the amount of work done is insufficient.

Playbot (Fig. 1) is designed to replace part of this control loop [13]. The user's visual system is still needed to determine the goal of a manipulation and to communicate with the robot. But the robot's visual system then takes its place in the closed-loop control of the robot in the execution of the task. The user is thus spared the frustration, tedium and effort of performing these tasks.

Let us imagine the following: A child is seated in a mobile, computer controlled wheelchair, which possesses a robotic arm with a manipulator, a stereo-colour camera system and a communication panel. The child would be able to point to an icon of a toy on the panel and then point to a sequence of action icons that he/she wishes the robot to perform with that toy, creating a sentence describing a play sequence. The play sequence could involve bringing toys to the child's table for close inspection and manipulation, for example. The wheelchair would visually locate the toys in the environment, plan the execution of play, and together with the child move and carry out the actions. The project is currently beginning its second phase, that is, implementation on a motorized wheelchair.

Playbot consists of a modified electric wheelchair (the Chair-man Entra, by Permobil Inc., USA), a 6+2 d.o.f. robotic manipulator (MANUS, by Exact Dynamics, Netherlands), a tablet PC, a number of monocular and binocular cameras, control electronics, three on-board laptops (Apple MacBooks with Intel Core 2 Duo Processors and 1GB RAM) running Linux and an off-board server (Sun Fire X2100 with a dual core AMD Opteron 1.75, 4GB RAM) also running Linux and providing further computational power. Both the electric wheelchair and manipulator were selected for the project due to their widespread clinical use. Further modifications to the wheelchair involved the integration of a motion controller by RoboteQ Inc. and the development of custom control electronics using a Motorola HCS12 microcontroller.

The co-ordination of services is handled by DataHub, a computing system that our Playbot group has developed. DataHub is based on a publish/subscribe model with discoverable services. Each service publishes information about itself in a central repository called Hub. When a client wants to subscribe to a service, it requests information from Hub about the service (IP address, TCP port and other related information) and Hub replies with the desired information. From that point on the client can talk to the service directly. In our approach, communication can continue even if Hub fails. Similar features are found in popular robot control software such as Player/Stage [4] and YARP [18].

The user interface is based on the BLISS symbolic language and is composed of various pictorial symbols representing locations, objects and actions. The operator can compose commands by selecting a sequence of iconic symbols on the tablet PC, such as "go to", "toy table", "point to", "target one", which tells the wheelchair to go to the Toy Table, to visually search for Target One and once Target One is detected, to raise the arm and point towards it. Control for Playbot is based on behaviors that combine deliberative and reactive processing with vision as the primary sensor [3], [10], [11].

III. A Description of the Related Behaviors

A. Door Localization

Standard template correlation/least squares based methods tend to degrade in performance when trying to localize specular objects from close distances, such as door handles, as the specularities become much more pronounced. Furthermore, depth extraction algorithms perform quite poorly when dealing with specular objects, making pure stereo vision based approaches problematic. The door localization and opening behaviors provide methods for dealing with these problems – using exclusively vision based sensors – and subsequently opening the door. This allows us to segment the door handle and thus know the exact region where we are going to grasp the handle. This is also vital for our depth extraction algorithm.

The handle detection of the door opening behavior has to be extremely reliable – low number of false positives so as not to frustrate the user – and capable of dealing with specularities and changes in illumination. Recognition methods based on edges/lines/image derivatives are known to be more robust than appearance based techniques under illumination and contrast changes. To achieve some robustness under changes in contrast and illumination, we use an algorithm introduced by Viola and Jones [16], which is based on a set of Haar-like features. As training data for each door handle, it uses a dataset of 1500 instances of the door handle at slightly different rotations and with different artificially induced illumination and contrast conditions. All the 1500 instances of the door handle were artificially generated using a single template image of the door handle.

One or more candidate handles are typically detected and we, therefore, need a method to remove any potential false positives. A simple template matching approach in HSV space is used. The template of the door handle and each detected region is converted to HSV space and a histogram comparison in the S channel spaces provides us with the candidate door handle. The candidate region most closely matching the door handle's template is selected as the segmented door handle region (Fig. 2(c)). To compare the S channels of two image regions, we first normalize each region's S channel histogram to have a total area of 1. The final matching score for the

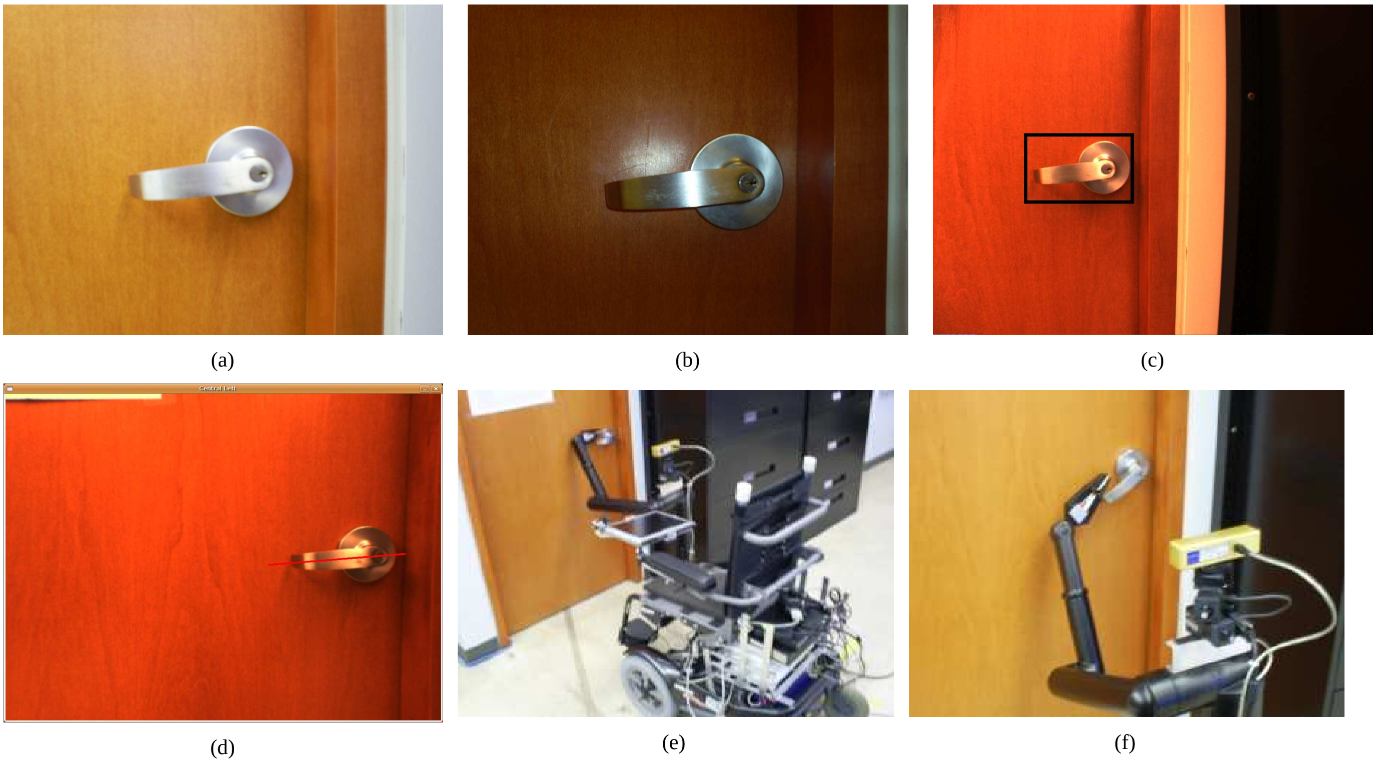

Fig. 2. (a): A close up view of a door handle. (b): The same door handle under different lighting conditions. Notice the intense specularities which can easily confuse recognition algorithms and stereo vision based depth extraction algorithms. (c): A detected door handle. Notice the slight difference in the door handle geometry compared to handles (a) and (b). (d): The extracted medial line which intersects the keyslot region where depth extraction is reliable. (e): Playbot's arm as it begins to open the door handle. (f): The arm a few frames later as it is in the process of opening the door by pushing the handle downwards.

two image regions is the total area of intersection of the two histograms. We have also performed histogram matching using the HS channels rather than only the S channel, but we observed a modest decrease in the door handle detection performance. We attribute this to the fact that the door handle contains white/grey/black, sometimes making the H channel histogram ill-conditioned.[1]

B. Door Opening

We extract reliable depth values from any regions of the door handle where conspicuous features exist. In our case this region corresponds to the keyslot location, as it is seen in Fig. 2(a)-(d). We use the relative coordinate system implied by the segmented door handle region and the known geometry of the door handle to accomplish this. A Canny edge detector is used in conjunction with a probabilistic Hough transform in order to detect the upper and lower parts of the door handle falling in the segmented door handle region. The parts are detected by searching in the segmented door handle region for the two most horizontally parallel lines that have a certain minimum amount of separation. As the lines detected by the Hough transform do not always detect the entire length of the upper and lower region of the door handle we extend the detected lines to cover the entire segmented region. From those two lines we extract their medial line which also intersects the lock region, as Fig. 2(d) shows. We use this medial line to detect the keyslot location by searching for the darkest region close to the right-half of the line. A precondition of our algorithm is that the stereo camera is almost parallel to the door. We extract 25 3D coordinates corresponding to 25 pixels around the keyslot region and fit those points to a plane using a least squares approach. This plane allows us to extract the orientation of the door handle with respect to the stereo camera coordinate system. If the plane angle/orientation along the axes is not within some reasonable limit ($\pm 10^\circ$) we assume we are dealing with poor depth extraction and set the plane parallel to the stereo camera. We also experimented without using the handle's orientation extraction algorithm, by assuming that the stereo camera was perfectly parallel to the door. This did not appear to significantly affect the performance of the algorithm, since the wheelchair's stereo camera is almost always placed parallel to the door. However, in future work, to achieve handle localization and manipulation from multiple aspects, estimation of handle orientation will become a necessity.

The 6 d.o.f. arm is then used to open the door. A single contact point on the door handle is defined and the arm is used to push that contact point along a desired circular trajectory followed by a forward push on the door itself, that pushes the door open. This single contact point simplifies the task and makes the solution more reliable as no grasping is needed and we only need to solve the inverse kinematics problem in order

[1]We thank one of the anonymous reviewers for making this suggestion.

to push the contact point along the desired circular trajectory without having to worry about force/torque control of the arm. However, force/torque control with compliance would likely simplify the problem and lessen the need for very accurate vision modules. As our robot arm is not equipped with any force sensors, all arm control is purely vision based.

A constrained optimization problem is defined and solved that controls the 6 parameters of the arm joints under the constraints that the gripper is maintained parallel to the door at all times and it is performing the desired circular motion to turn the handle. Once the handle is turned, a forward push on the door's body can open the door. A video of the arm opening behavior can be found at [13].

IV. Experimental Results

Playbot (Fig. 1) is positioned near the door handle at various distances (stereo camera distances of approximately 45cm-72cm) from the door handle, and at a number of modest rotations from its default parallel position to the door (within approximately $\pm 10°$), in order to evaluate the door localization and door opening behavior's performance. We made an effort to position Playbot at a number of poses that would approximately uniformly sample the space of distances (45cm-72cm) and the space of orientations ($\pm 10°$). All sensing is done using a stereo camera (Bumblebee stereo camera, Point Grey Research Inc.). We executed the behaviors 15 times with a success rate of $12/15$ runs. Our results are presented in Fig. 3. Two of the failed cases occured because the stereo camera was too close to the door and as a result the door handle did not fall within the field of view of both the left and right camera views. The other failed case occured at the other extreme of distances (72cm). At this distance the arm had to be almost fully extended to reach the handle. Poor depth extraction and/or poor optimization in the inverse kinematics caused the estimated joint angles to fall in local minima. However, the door handle and the keyslot were still accurately located at this distance. The results demonstrate the importance of using cameras with wide fields of view or incorporating some mechanism for automatically adjusting the wheelchair's distance from the door, so that the door handle always falls within the field of view of the cameras.

Without our improved methodology for depth extraction and due to the specular nature of the door handle, the system typically over-estimates the depth of a random pixel on the door handle. On 15 stereo images acquired during our test runs, the estimated depth at any pixel on the door handle – other than the keyslot location – was typically 8-10cm greater than its true depth. This occured due to the specular nature of the reflection on the door handle and the small number of conspicuous features on the handle's surface. Errors of 8-10cm in the depth extraction can cause the arm to collide with the door and our module to fail. In all cases, we used the Birchfield-Tomasi algorithm to determine the stereo correspondences [1].

There was only a modest amount of variability in the illumination, as all our experiments were done indoors and in a controlled environment. Tests were performed during various moments of the day (morning, evening, etc...) which affected the illumination conditions in the lab –large windows are located in the lab–, however, this did not appear to significantly affect the algorithm. However, we performed an experiment with an intense source of light falling on the door handle, which caused the histogram comparison to be less accurate, indicating that more work can be done to improve the robustness of the algorithm in this respect.

V. Conclusion

An approach for localizing and opening doors using a robot wheelchair equipped with a 6+2 d.o.f. robotic arm was presented. All sensing was exclusively vision based. Our results demonstrate the reliability of our approach. We are currently in the process of extending this door handle localization algorithm by making the system capable of detecting door handles that do not fall within our current field of view by actively controlling the wheelchair orientation and the camera's pan/tilt. The problem of searching for an arbitrary object in 3D space is NP-hard and it is, therefore, intractable [20]. In [12], [15], [19] an active approach to visual search which uses a number of heuristics to circumvent the inherent intractability of the problem was presented. The active approach for visual search is necessary for a number of reasons:

- To be able to move the fixation point/plane or to track motion
- To see a portion of the visual field otherwise hidden due to occlusion (manipulation, viewpoint change)
- To see a larger portion of the surrounding visual world (exploration)
- To compensate for spatial non-uniformity of a processing mechanism (foveation)
- To increase spatial resolution or to focus (sensor zoom, observer motion, adjust camera depth of field, stereo vergence)
- To disambiguate or to eliminate degenerate views
- To achieve a pathognomonic view [17].

We are currently in the process of incorporating this work into Playbot. Our current implementation works on a single type of door handle. Future work involves making the system capable of handling arbitrary door handles, by automatically detecting the most conspicuous features that should be used for reliable stereo depth extraction from an arbitrary new door handle, and being capable of learning the appropriate arm motion that would open an arbitrary door. We are investigating tactile methods based on force feedback to accomplish this. Finally, future work involves making the system capable of opening a door by pushing the door handle and pulling – rather than pushing the door handle and pushing the door as the current system does – and also making the system capable of closing the door. Service robots can help people with mobility impairments lead a more independent and productive life. The work presented here falls within this context, demonstrating that the confluence of robotics and advanced vision algorithms is promising and important for all.

distance (cm)	angle (deg)	Handle located	Keyslot located	Door opened
45	12	No	No	No
45	-10	No	No	No
46	0	Yes	Yes	Yes
47	0	Yes	Yes	Yes
50	7	Yes	Yes	Yes
51	0	Yes	Yes	Yes
52	-7	Yes	Yes	Yes
56	0	Yes	Yes	Yes
60	6	Yes	Yes	Yes
63	0	Yes	Yes	Yes
63	-7	Yes	Yes	Yes
66	7	Yes	Yes	Yes
68	0	Yes	Yes	Yes
68	-6	Yes	Yes	Yes
72	7	Yes	Yes	No

Fig. 3. Our results for various distances of the stereo camera from the door and various modest rotations of the camera. An angle of zero corresponds to the stereo camera being parallel to the door, and a positive angle corresponds to a counterclockwise rotation of the camera when viewed from above. All values are rounded to the nearest integer. As it is seen, problems arise when the camera is too close to the door ($<$46cm) as typically then the door handle is not in the field of view of both the left and right views of the camera. Rotations of the camera near the door make it more likely that this problem will occur. Another error occured at a large distance (72cm). At this depth the arm had to be fully extended. We attribute the error to poor depth extraction and/or to the optimization algorithms used in the inverse kinematics estimation falling in a local minimum. Otherwise, the algorithm is fairly reliable.

ACKNOWLEDGMENTS

JKT holds the Canada Research Chair in Computational Vision and gratefully acknowledges its financial support. AA gratefully acknowledges the financial support of the NSERC PGS-D program. We also thank the reviewers for their insightful comments and suggestions.

REFERENCES

[1] S. Birchfield and C. Tomasi "Depth Discontinuities by Pixel-to-Pixel Stereo", International Journal of Computer Vision, 35(3): 269-293, December 1999.

[2] R. Brooks , L. Aryananda, A. Edsinger, P. Fitzpatrick, C. Kemp, U. O'Reilly, E. Torres-Jara, P. Varshavskaya, J. Weber "Sensing and manipulating built-for-human environments", International Journal of Humanoid Robotics, 1:1, pp.1-28, 2004.

[3] S. Dickinson, S. Stevenson, E. Amdur, J. Tsotsos, L. Olsson,"Integrating Task-Directed Planning with Reactive Object Recognition", SPIE Conference, Boston, Sept. 1993

[4] B. Gerkey, R. T. Vaughan and A. Howard. "The Player/Stage Project: Tools for Multi-Robot and Distributed Sensor Systems". In Proceedings of the 11th International Conference on Advanced Robotics (ICAR 2003), pages 317-323, Coimbra, Portugal, June 2003.

[5] O. Khatib "Mobile manipulation: The robotic assistant" Journal of Robotics and Autonomous Systems, Vol. 26, pp175-183, 1999.

[6] K. Nagatani, S. Yuta. "Designing Strategy and Implementation of Mobile Manipulator Control System for Opening Door". Proceedings of the IEEE International Conference on Robotics and Automation, Minneapolis Minnesota, April 1996.

[7] G. Niemeyer, J.J.E. Slotine. "A simple strategy for opening an unknown door", In Proc. of the 1997 IEEE International Conference on Robotics and Automation. Albuquerque, New Mexico. April.

[8] L. Peterson, D. Austin, D. Kragic. "High Level Control of a Mobile Manipulator for Door Opening", In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, 2000.

[9] C. Rhee, W. Chung, M. Kim, Y. Shim, H. Lee. "Door opening control using the multi-fingered robotic hand for the indoor service robot", In Proceedings of the IEEE International Conference on Robotics and Automation, New Orleans LA, April 2004.

[10] A. Rotenstein, A. Rothenstein, J.K. Tsotsos, "Directed Behaviour-Based Robots For Planetary Exploration", Proc. i-SAIRAS'03, Nara, Japan, May 2003.

[11] A. Rotenstein, A. Rothenstein, M. Robinson, J.K. Tsotsos, "Robot middleware should support task-directed perception, too", 2nd International Workshop on Software Development and Integration into Robotics, International Conference on Robotics and Automation, Rome, Italy, April 14 , 2007.

[12] K. Shubina. "Sensor Planning for 3D Object Search", M.Sc. Thesis, Department of Computer Science and Engineering, York University, 2007

[13] The Playbot Project. URL: `http://www.cse.yorku.ca/~playbot`

[14] J. Tsotsos, G. Verghese, S. Dickinson, M. Jenkin, A. Jepson, E. Milios, F. Nuflo, S. Stevenson, M. Black, D. Metaxas, S. Culhane, Y. Ye, R. Mann. "PLAYBOT: A visually-guided robot to assist physically disabled children in play", Image and Vision Computing, Special Issue on Vision for the Disabled" p275 - 292, April 1998.

[15] J.K. Tsotsos, K. Shubina. "Attention and Visual Search: Robotic Vision Systems that Search", IEEE International Conference on Computer Vision Systems, 2007.

[16] P. Viola, M. Jones. "Rapid Object Detection using a Boosted Cascade of Simple Features", IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 1, p.511, 2001.

[17] Wilkes, D., Tsotsos, J.K., "Active Object Recognition", CVPR-92, Urbana, Ill, 1992 , p 136 - 141

[18] YARP project. URL: `http://eris.liralab.it/yarp/`

[19] Y. Ye, J.K. Tsotsos, "Sensor Planning for Object Search", Computer Vision and Image Understanding, 73-2, p145-168, 1999.

[20] Y. Ye, J.K. Tsotsos,"A Complexity Level Analysis of the Sensor Planning Task for Object Search", Computational Intelligence Vol. 17, No. 4, p. 605-620, Nov. 2001.