TOWARDS THE DREAM OF AN INTELLIGENT, VISUALLY-GUIDED WHEELCHAIR

Andrei M. Rotenstein, Alexander Andreopoulos, Ehzan Fazl, David Jacob, Matt Robinson, Ksenia Shubina, Yuliang Zhu, and John K. Tsotsos
Department of Computer Science and Engineering, and Centre for Vision Research, York University

INTRODUCTION

An important focus of research into smart wheelchairs aims to assist people living with physical disabilities and possibly without speech but with normal cognitive functioning, vision and audition. What is sought is a wheelchair that is *intelligent*, that can visually understand the world and will do precisely what its operator wants, without tedious and detailed user control. This dream has yet to be achieved [21].

These potential wheelchair users have the ability to issue high-level, sentential commands resembling natural language. Commands may include references to objects present in the environment that can be observed visually by both the smart wheelchair and its user. Given this, a truly intelligent wheelchair robot should be able to interpret the commands accurately and execute them safely. Required capabilities include object recognition, knowledge of objects and locations, landmark-based navigation, visual search, planning and reasoning about tasks, executive control and failure recovery. Also required is a suitable user interface and mode of interaction. For example, a user might command the robot to open a door to a room, go to a table and pick up a few toys on top of it. The smart wheelchair must detect the door handle, grasp it accurately it so that it can be turned, maneuver so that the door is opened, navigate the doorway (often only a few centimeters wider than the wheelchair), identify the table location and navigate the room towards it—avoiding people or other obstacles—visually identify the objects to be picked up and, finally, grasp and manipulate the objects. It must do so safely and without unnecessary motion, while monitoring for failures and recover from them sensibly, without excessive user intervention.

While there has been much headway on specific issues, such as robot control architecture, artificial reasoning, simultaneous localization and mapping, manipulator control and object recognition, the research community has achieved little towards integrating these into the system envisioned.

To address this, we have developed a novel robotic wheelchair platform, PlayBot, for conducting basic research into robot control, computer vision and automated reasoning for visually-guided smart wheelchairs. The wheelchair, complete with highly dexterous robotic manipulator, is equipped with a number of monocular and binocular cameras. Only vision is required, since both the user and the wheelchair can visually perceive the same environment. The user employs a pictorial/iconic interface to specify high-level command sentences. Currently, our platform performs map-based navigation, visual collision avoidance, and can search for and point the arm toward toys on tables. We employ a behaviour-based approach to control, which is built on top of a custom distributed computing system.

This paper describes the design PlayBot. We begin with the project's motivation and goals in the Background, and continue with the hardware, software and user interface in the PlayBot System section. We then discuss our current research activity as well as related work.

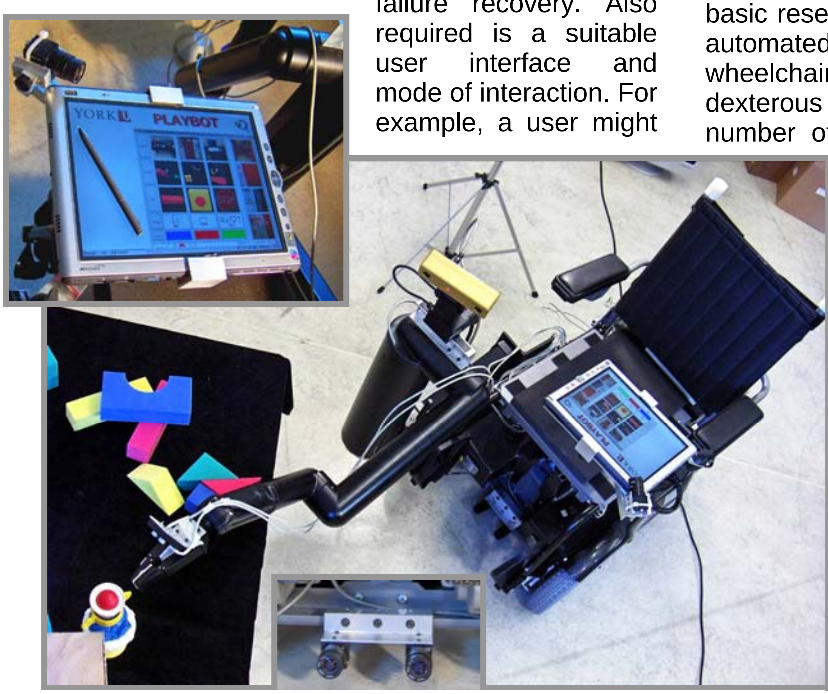

Figure 1: Playbot gesturing towards a toy on a table top.

BACKGROUND

The PlayBot project started in the early 1990's [23] to address limitations of assistive robotics of the time. Users controlled such systems through a long series of microactivations, which effectively relied on the user's vision to be an integral part of a closed-loop control system. This could be tedious, frustrating and, consequently, tiring; however it afforded the user some personal freedom. The motivating notion behind PlayBot is that a vision-based robot control system could replace the function of the user's visual system, in order to reduce the user's fatigue and frustration, while providing the same personal freedom. This idea continues to motivate the project today.

PlayBot's goals include: i) to *use vision* as the key sensing modality, as humans do; ii) for perceptual processing be *task-directed*, as it has been long established that task direction is a key strategy in reducing its complexity; iii) *visual search*, which we believe to be a critical precondition of most tasks that a smart wheelchair must perform; iv) *safety and robustness in non-engineered environments*, to minimize the need for caregiver intervention. Many important contributions were made towards these goals, during the early years of this project, including visual attention with a stereo head [22], gaze stabilization [8], 3d object search [26] and active object recognition [25]. However, until now, PlayBot did not exist as an integrated smart wheelchair platform.

THE PLAYBOT SYSTEM

Platform

PlayBot consists of a modified electric wheelchair (the Chair-man Entra, by Permobil Inc., USA), a differentially-steered vehicle with a chair that lifts, a 5+2 d.o.f. robotic manipulator (MANUS, by Exact Dynamics, Netherlands), a Windows-based tablet PC, a suite of monocular and binocular cameras, a number of on-board Linux-based computers and custom control electronics. Both the electric wheelchair and manipulator enjoy widespread clinical use, which is why these devices were selected for this project.

The system is controlled by a multi-level control architecture. The top level is a novel intelligent, behavior-based control architecture [23], [16] with an execution system, and is suited for development by a team of designers [17]. A middle level consists of a Motorola HCS12 microcontroller, while the lowest level consists of off-the-shelf devices for controlling the wheelchair motors and electronic arm. The arm has a two-finger gripper and is mounted on a two-position lift. Pieper's solution [13] is applied in computing the inverse kinematics of the arm, given a desired pose of the end-effector. To avoid collision with the wheelchair, we limit joint angles, and we define atomic motions, such as folding, unfolding, pointing and grasping, that are known to not intersect with the wheelchair.

PlayBot is pictured in Figure 1, extending its arm towards a toy on top of a table. Gazing at the toy from its perch on the arm's shoulder joint is a Point Grey Bumblebee camera mounted on a Directed Perception Pan/Tilt Unit. Two other stereo cameras are also shown here: one is a pair of Point Grey Flea cameras, mounted facing forward at the base of the wheelchair, in between the wheels (also inset, bottom-centre); the other is a pair of webcams mounted on a metal bracket on the arm's end-effector (two USB wires can be seen in the picture above feeding into the camera pair). Another Point Grey Flea camera is shown above mounted on the tablet PC (also inset, top-left). The tablet PC presents the wheelchair user with a graphical display of images and Bliss language symbols (also inset, bottom-right).

Construction of the wheelchair robot involved extensive modifications to the electric wheelchair, including integration of a DC motor controller by RoboteQ, Inc., and the development of custom control electronics based on a Motorola HCS12 microcontroller, including a switching power supply that has software control. Also employed are three on-board laptop PCs and a number of off-board Linux-based computers.

User interface

The goal of the PlayBot user interface is to provide a means for the disabled user to accomplish complex tasks with a minimum of mental and physical effort. In our interface, a non-technical user is able to express a high-level command with the form "*verb* {*modifier*} *object*" [24], using a pictorial/iconic interface inspired by the Bliss symbolic language, invented by Charles Bliss in 1941, and used by cerebral palsy patients to communicate [7]. The interface presents the available objects, locations and actions to the user. The user can point to a pictorially represented action and then to a picture of an object, and thus compose a grammatically formed sentence, such as, "GOTO TABLE, POINT RED_BLOCK", which PlayBot then translates into a plan for execution. The interface is shown in Figure 2.

Software System

The software system consists of an intelligent control architecture that is build on top of a distributed computing platform. The control architecture consists of an extensible, distributed behaviour-based system that allows for deliberative processing to be integrated into behaviours. It is coordinated by an iterative-repair based execution system.

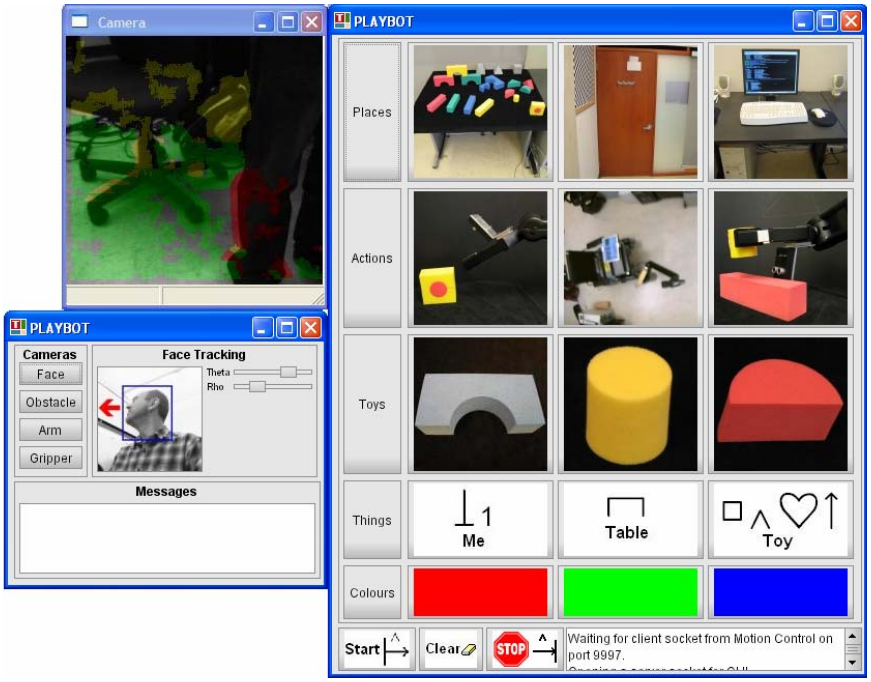

Figure 2: PlayBot's pictorial/iconic user interface.

Distributed computing is handled by DataHub, our simple cross-platform system for coordinating services. Services are processes that handle specialized tasks, such as controlling the wheelchair motors, arms, or do visual sensing activities. A service accepts requests to perform operations from processes that are running on different hosts and reply with the results. DataHub is cross-platform, supports multiple programming languages, such as C/C++, Java, and Tcl, and it operates over TCP/IP.

DataHub is based on a publish/subscribe model with services that are discoverable. A service publishes interface information in a repository service called *Hub* that resides on another host. The Hub sends service connection information to subscribing clients, which can then talk to the service directly. This differentiates DataHub from other approaches, such as CMU's InterProcess Communication (IPC) [20], which has a process responsible for handling all communication between services and clients and is ultimately a performance bottleneck. In DataHub, communication can continue even if Hub fails.

To execute this high-level command, the executive co-ordinates behaviours to carry out each atomic task in the operator's sentence. The executive uses an iterative repair algorithm to refine the user's command sentence into atomic steps of a plan. The atomic steps are translated into a behaviour network, which executes each task in a reactive manner. Behaviours include: *Drive*, for controlling the platform; *Detect*, for detecting obstacles; *Avoid*, for navigating obstacles; *Reach* for performing manipulator actions; and *Search*, for performing a view-based object search strategy [19], [26].

RESEARCH THRUSTS

Currently, our system is capable of accepting commands to navigate towards known locations in our lab environment and to visually locate and point to known objects on tables, such as "GOTO *DESK1*; POINT *BLOCK1*". It employs a static map of our laboratory. This is the launching point of our research activity.

Vision is our main focus. For example, we are constructing a network of intelligent overhead cameras for assisting with navigation and localization and for monitoring the wheelchair's environment, especially where the wheelchair cannot see. It will be installed throughout a corridor and several adjoining rooms of our environment, effectively making the environment "smart". We do not want to rely on a highly engineered environment, however. To this end, we are examining approaches to landmark-based navigation with learning. Also under investigation are object recognition, visually guided reaching and grasping, and visual obstacle avoidance; our focus here is on employing task-directed processing to solve some of these problems. Since it is the intelligent control software that integrates these technologies, we are developing better approaches to executive control of behaviour-based controllers that employ task-directed visual perception extensively.

In addition, we require further investigation into user interface issues. Currently, we employ a tablet PC, however it must be improved to support *task transparency* [18]. That is, the user should be able to interact directly at the level of the robot task, instead of indirectly, via operational tasks (interaction with the wheelchair itself). This could be achieved if the interface only presented symbols relevant to what robot tasks are possible in the current context, and can adjust this as the context changes. Usability issues also need to be investigated.

RELATED WORK

Current smart wheelchair research projects range in their applications, capabilities and use of sensors. Some projects focus on users with cognitive impairments. The simplest typically employ bumpers and sonar for collision avoidance and sometimes follow lines [14], [11], and select a safe driving direction using joystick input and sonar readings, but often cannot navigate doorways. Some achieve better navigation by employing a laser rangefinder and IR sensors for obstacle-avoidance, wall-following and three-point turns [15]. Many approaches specialize in safe navigation of highly dynamic environments, such as subway stations [10]. Some systems perform landmark-based navigation and employ visual behaviours for greater autonomy (e.g. [6]), but do not provide a manipulator. In contrast, PlayBot's goal is to perform visual tasks at a higher level of abstraction.

Current approaches employ a variety of interface methods, such as touch-screen or voice recognition interfaces [4], [1], [5]. Others detect eye movements via eye trackers either to guide the wheelchair directly with the direction of gaze [9], or to select menu items on a display [3]. Facial expressions, mimic and gesture recognition have also been used to as input [9], [3], [4]. Additionally, "sip and puff" devices and single switches that are common in rehabilitation technology have also been used for smart wheelchairs [12]. While our project employs a pictorial/iconic interface, PlayBot is not restricted to this: it is conceivable that a user could also construct sentences via such interface devices; for example, by composing sequences of saccades or sips and puffs.

ACKNOWLEDGEMENTS

We gratefully acknowledge the financial support of Precarn, Inc., the Canada Research Chairs programme, the Canadian Foundation for Innovation, NSERC and York University.

REFERENCES

[1] A. Argyros, P. Georgiadis, P. Trahanias, and D. Tsakiris. "Semi-autonomous navigation of a robotic wheelchair." *Journal of Intelligent and Robotic Systems*, 34:315–329, 2002.

[2] Z. Bien and W.-K. Song. "Blend of soft computing techniques for effective human-machine interaction in service robotic systems." *Fuzzy Sets and Systems*, 134:2–25, 2003.

[3] Z. Bien, M.-J. Chung, P.-H. Chang, and D.-S. K. et al. "Integration of a rehabilitation robotic system (KARES II) with human-friendly man-machine interaction units." *Autonomous Robots*, 16:165–191, 2004.

[4] F. Bley, M. Rous, U. Canzler, and K.-F. Kraiss. "Supervised navigation and manipulation for impaired wheelchair users". In W. Thissen, P. Wierings, M. Pantic, and M. Ludema, eds., *Proc. IEEE International Conference on Systems, Man and Cybernetics: Impacts of Emerging Cybernetics and Human-Machine Systems*, pp 2790–2796. The Hague, IEEE Systems, Man& Cybernetics Society, October 2004.

[5] J. D. Crisman, M. E. Cleary, and J. C. Rojas. "The deictically controlled wheelchair." *Image and Vision Computing*, 16(4):235–249, March 1998.

[6] T. Gomi and K. Ide. "The Development of an Intelligent Wheelchair." *Proceedings of IVS'96*, Tokyo, Japan, 2006.

[7] E. Helfman, *Blissymbolics: Speaking without Speech*. Elsevier/Nelson Books, New York, 1981.

[8] A. Jepson, M. Black. "Mixture models for optic flow computation", In I. Cox *et al*., eds*., Proc. DIMACS Workshop on Partitioning Data Sets*, AMS, Providence RI, 1995.

[9] Y. Kuno, N. Shimada, and Y. Shirai. "Look where you're going: A robotic wheelchair based on the integration of human and environmental observations." *IEEE Robotics and Automation*, 10(1):26–34, 2003.

[10] A. Lankenau, O. Meyer and B. Krieg-Bruckner. "Safety in Robotics: the Bremen Autonomous Wheelchair." In *Proc. of the 5th Int. Workshop on Advanced Motion Control*, 1998.

[11] Levine, S.P., Bell, D.A., Jaros, L.A., Simpson, R.C., Koren, Y., and Borenstein, J. "The NavChair Assistive Wheelchair Navigation System." *IEEE Transactions on Rehabilitation Engineering*, Vol. 7, N0. 4, December 1999, pp. 443-451.

[12] M. Mazo *et al*. "An integral system for assisted mobility." *IEEE Robotics and Automation Magazine*, pp. 46–56, March 2001.

[13] D. Pieper, and B. Roth. "The kinematics of manipulators under computer control", *Proc. 2nd International Congress on the Theory of Machines and Mechanisms* (Vol, 2), Zakopane, PL. pp159—169, 1969.

[14] Nesbit, P.D. "Assessment and training of Children for Powered Mobility in the UK." *Technology and Disability* 14, pp. 173—182, Ios Press, 2002.

[15] R. S. Rao, K. Conn, S. H. Jung, J. Katupitiya, T. Kientz, V. Kumar, J. Ostrowski, S. Patel, C. J. Taylor. Human Robot Interaction: Application to Smart Wheelchairs. *IEEE International Conference on Robotics and Automation 2002*, Washington, D.C.

[16] Rotenstein, A.M. "Supporting deliberation within behaviour-based systems", M.Sc. Thesis, Department of Computer Science, York University, 2003.

[17] Rotenstein, A. M., Rothenstein, A. L., Robinson, M., Tsotsos, J.K. "Robot middleware must support task directed perception*", ICRA Workshop on Software Development and Integration in Robotics*, 2007. (To appear.)

[18] Shein, G. F., "Towards Task Transparency in Alternative Computer Access: Selections through Switch-Based Scanning", PhD Thesis, University of Toronto, Canada, 1997.

[19] K. Shubina, "Sensor planning in 3D object search", M.Sc. Thesis, Department of Computer Science and Engineering, York University, 2007.

[20] R. Simmons, D, Apfelbaum. "A task-description language for robot control", *Proc. IROS,* Victoria, BC., 1998.

[21] R. C. Simpson, "Smart Wheelchairs: A literature review." *J. Rehabilitation Research and Development 42* (4), pp423-438, 2005.

[22] J.K. Tsotsos, S. Culhane, W. Wai, Y. Lai, N. Davis, and F. Nuflo. "Modeling visual attention via selective tuning", *Artificial Intelligence 78* (1-2), pp. 507—547, 1995.

[23] Tsotsos, J.K., "Intelligent Control for Perceptually Attentive Agents: The S* Proposal", Robotics and Autonomous Systems 21-1, p5-21, July 1997.

[24] J.K. 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 for physically disabled children", *Image and Vision Computing 16,* pp 275—292, 1998.

[25] D. Wilkes, J. K. Tsotsos, "Integration of camera motion behaviours for active object recognition", in *Proc. IAPR Workshop on Visual Behaviours,* Seattle, 1994.

[26] Y. Ye, "Sensor planning in 3D object search", Ph.D. Thesis, Department of Computer Science, University of Toronto, 1997.