DEVELOPMENT OF A ROBOTIC COMPANION: A SURVEY

Anna Maria Joseph[1], M Mathurakani[2], Nikhil Malhotra[3]

PG Scholar, Department of Electronics and Communication, TocH Institute of Science & Technology, Ernakulam, India, mariaannnajoseph@gmail.com[1]

Prof. Head of department (M Tech), Department of Electronics and Communication, TocH Institute of Science & Technology, Ernakulam, India, Formerly Scientist , DRDO, Ministry of Defence India[2]

Principal Technical Architect, Tech Mahindra Ltd, Pune, Maharashtra, India[3]

ABSTRACT- Making life easier, safer and efficient has been always a constant thrust of innovative minds. From healthcare and homecare, to military use and emergency response, robots are fast becoming a fixture in our lives. When it comes to the development of an autonomous robot, the first step is to study the methods for developing one. A robot consists of sensors to perceive the environment and the own movement, on-board computers, and actuators for vehicle control. The Kinect sensor is an advanced computer vision component with a variety of useful features that makes it suitable as the vision system for an autonomous robot. In this paper, the approach taken for the development of the autonomous robot is discussed.

Keywords: Robotics, 3D Sensing, Depth vision, Autonomous companion, Kinect, Human Recognition, neural network

INTRODUCTION

Innovative ideas along with sophisticated new technologies has been always a boost to mankind in developing new and improved systems to reduce human effort in day to day life making life easier at every point. At our present world, Shopping centers have grown up into one of the busiest business points where anything and everything is available. People gather up at such easy shop away spots for convenience but most of the time it is observed customers struggles to carry around their chosen products and have to face long queues just to get billed for the products chosen. It is in one such situation, an innovative mind thought about a robot to assist the customer in shopping. The thought, along with the desire to realize it was the starting point of this project**.** His quest for implementing the project led to Microsoft Kinect which was the stepping stone to Autonomous Robotic companion.

A robotic companion, with ability of user, voice and gesture recognition, capable of accompanying you wherever you move, avoiding all obstacles that hinders its path. It sounds like a science fiction movie, but that is exactly what our project aims to achieve. When I say user recognition, the robot should be able to identify you when you come in front of the robot. And it should be able to recognize you as the primary user. The ability to understand what you are saying and replying back accordingly will make the robot to be a friend to you when you are alone at home. With the wave of your hand or with a voice command, the robot should follow your instructions. When guests are at home, a robot assisting you to carry the tray for you, instead of you going back and forth will surely catch the attention. When you want to shop online, it would add to your confidence if you can see exactly how the dress will look on you. The robot can act as a virtual trial room for you. Or may be in a retail shop, you need not wait for the trial room to be free. The robotic companion can be your shopping companion to carry the items when you are going for your routine grocery shopping.

An autonomous robotic companion is a good solution for a whole class of problems for the elderly too: a robotic butler that helps on carrying heavy objects; a robotic assistant to remind the time to take medicine; an automatic walking aid that should support elderly people and so on.

LITERATURE SURVEY

For developing an autonomous robotic system, the sensory perception of the robot and the kinematics for the robot movement are two important factors. The environment, in which the robot moves, has to be perceived accurately for the efficient utilization of the robot and to avoid any mishap. A robot which can be used as a companion for a human requires the robot to identify the person and interact with him. A sensory system capable of doing this efficiently and the movement of the robot is the focus of this study. Microsoft Kinect, launched in November 2011 for the Xbox 360 gaming console, is a motion sensing input device capable of tracking the user and takes audio data also.

As the first step to develop a robot, some of the recent autonomous robots were studied

ROBOTS IN HUMAN LIFE

'Rachkham' developed by Clodic et al [1] is an interactive robotic tour guide developed based on visual human robot interaction. Kirby et al [2] designed 'Grace' in 2007 to accompany a person side-by-side and engage him like a human would do, while maintaining a minimum space between the human and robot. A laser based tracking method was implemented in using direction following and path following. If the user goes out of the sensor range, it was informed by the robot. But it was found that the robot responds only to the person's speed and location, ignoring aspects such as the person's identity or personality, spoken or gestured commands from the person. 'Minerva' is tour guide robot developed by Thrun et al[3] to escort visitors in the Smithsonian Museum of American History. Johnny developed in 2012 by Breuer et al [4], based on the RoboCup@Home challenge [5], was designed to serve in a restaurant-like environment, where it received seat reservations, waited on guests, and delivered orders to them. Another well-known service robot is 'BIRON', developed in 2004 by Haasch et al. [6], and designed to actively interact with its user by means of natural user interface. MKR is an omni-directional mobile transfer robot system, developed in 2010 by Takahashi et al [7], for hospital applications. Using virtual potential field methods, the robot could transfer luggage, specimens and other important materials to its goal avoiding obstacles in the path.

The area of service robotics is broadening from robotic vacuums, bomb retrievers, exoskeletons and drones, to robots used in surgery, space exploration, agriculture and construction. A useful robotic system can be created by integrating various modules. Care-o-bot, a highly integrated and compact service robot**,** developed in 2009, by Reiser et al [8], is equipped with laser range scanners, a vision system and a 7 DOF manipulator arm. The sophisticated software for navigation, manipulation and vision, created a solid basis for the development of challenging manipulation tasks in everyday environments. As a typical application, the robot can act as a robotic butler, where a customer may ask for a drink using a touchscreen on the robot. The robot identifies the requested bottle using an object recognition module. The robot then lifts the bottle, places it in a tray and gets a cup from adjacent room to be placed along with the bottle to be served to the customer.

A robot can be used in domestic environments to serve in the everyday task or to assist the elderly. Taking care of elderly or a chronically ill person can be a challenge which can be simplified with a service robot. 'Flo', developed in 2000 by Roy et al. [9] is a service robot designed with a touch-sensitive display, a laser range finder, an array of 16 sonar sensors, and two on-board PCs. Its navigation system enables the robot to navigate safely in the indoor environment. It is equipped with, a telepresence software which would allow remote medical consultation, a speech interface for communicating with the user and neural network based face recognition system. In the present world, with a growing number of elderly people living alone, robots can be used to care this section of the population. 'Hector', developed by Gross et al. [10], under the 'CompaniAble' project, was designed to assist the elderly who suffer from mild cognitive impairment, in home environments. Hector can work collaboratively with a smart home and remote control center to better support older people living at home. As the field of socially assistive robots is maturing into useful technology, the next step is to study the different ways in which the task of making a robotic companion can be accomplished.

3D SENSING AND DEPTH RESOLUTION

For being a companion to a human, a robot must be capable of perceiving the environment as humans do. The sensory system of the robot must be capable of perceptual abilities like humans. In addition to standard capabilities like obstacle avoidance, and navigation, the robotic perceptions have to be optimized for interacting with humans on a human level. The robot should be equipped with sensors, to track human features, user's motion and at the same time detect any obstacle that hinders its path.

The complexity involved in human motion and its structural appearance makes the human motion analysis a challenging task. Aggarwal et al [11] gives an overview on the human motion analysis focusing on three major areas for motion analysis 1) motion analysis involving human body parts, 2) tracking of human motion using single or multiple cameras, and 3) recognizing human activities from image sequences. After successfully matching the moving human image from one frame to another in image sequences, understanding the human movements or activities comes naturally, which leads to a discussion of recognizing human activities.

Paola et al [12] uses a multisensory platform equipped with a monocular camera, a laser scanner, and an RFID device for autonomous mobile robot. Three layer architecture is used for the surveillance of the environment using a reconfigurable approach. Each of the three main components, i.e. controller, executor, and supervisor, is connected with the sensory input. This information is used in different ways: at the highest level, sensory data are converted into events, which are used to control task executions; at the middle level, sensory data are used to monitor and control the execution of the task in progress; finally, at the lowest level, sensory inputs are used by active behaviors to perform the associated actions. The limitation with this approach is for people and obstacle detection, the robot has to stop and stay there to process the data.

Visual camera can provide dense 3D information of a scene, but lacks depth information. Laser range finder can be used to provide accurate distance measurements from object in the environment. Maria et al [13] integrates a 2D laser range finder to obtain depth information from the environment. Main advantages of laser systems are a broad bandwidth and small beam divergence and footprint. They also offer a high immunity to atmospheric effects in opposition to the visual cameras. Laser range finder gives a sparse, but accurate map of the environment in a 2D plane. They use an infrared light beam of 905 nm, which receives directly the reflected signal from the objects, in polar coordinates. The laser operation mode is based on a time-of-flight (TOF) measurement principle: a single laser pulse is sent out and reflected by an object surface. The elapsed time between emission and reception allows the calculation of the distance between laser unit and object. The laser pulses sweep a radial range in front of the laser unit, via an integrated rotating mirror.

Chen and Birchfield [14] uses a stereo pair present the Binocular Sparse Feature Segmentation (BSFS) algorithm for vision-based person following. To determine the location of the person in the image, two images of a stereo pair as well as successive video frames are matched. The motion and stereo information are fused to handle difficulties like dynamic backgrounds; out of plane rotation etc. The system is subject to distraction by other objects with similar motion and disparity to the person being tracked.

When the robot moves in an environment populated with humans, following a leader while avoiding static and dynamic obstacles poses multiple challenges to the robot. Marchetti et al[15] describes how data fusion algorithm is utilized to combine both information, from the laser and the RGB-D camera to provide more reliable information to the robotic trajectory control system and ensure a smooth and precise execution of detection and follow of user overcoming all challenges in path.A data fusion algorithm has been developed using a particle filter, augmented with a k-clustering step to extract person estimations. The result on a real mobile platform verifies the effectiveness of the approach.

The advent of Microsoft Kinect, opened up new possibilities for human detection and tracking. Kinect was designed for natural interaction between the user and computer in the gaming environment. However, the characteristics of the data captured by Kinect have attracted the attention of researchers and developers in the field of retail, healthcare, medical, robotics etc.

KINECT

The Microsoft Kinect, with RGB camera, IR sensor and microphone array, is a revolutionary product that can sense human motion, location as well as their voices. These features plus the ability to control the user's console or computer with gesture or spoken commands opened up entirely new commercial experiences in multiple industries. Even though Kinect was developed for the Xbox 360 gaming console, drivers were developed for connecting it to a laptop. And Microsoft launched its own SDK for the Kinect in 2012. The Kinect v1 sensor can track skeleton of two people completely and 20 joints of each person. In October 2014, Microsoft launched the Kinect v2 sensor, with the ability to track as many as six people and 25 skeletal joints per person. The enhanced fidelity of the depth camera, combined with improvements in the software improved body tracking and the ability to see smaller objects.

Khoshelham [16] performs an accuracy analysis on the depth information obtained from the kinect sensor. Based on the mathematical model of depth measurement by the sensor a theoretical error analysis is presented, which provides an insight into the factors influencing the accuracy of the data. Compared to a laser data, a properly calibrated Kinect sensor has less systematic errors in the point cloud. With increasing distance from the sensor, the depth error increases quadratically and reaches 4cm for maximum range. The density of the points also decreases when the distance from the kinect sensor increases.

Shotton et al [17] proposes a new method to quickly and accurately predict 3D positions of body joints from a single depth image. This method implemented by Microsoft SDK for person detection and tracking, takes an object recognition approach, and runs at 200 frames per second. To detect the human body parts, it uses a deep randomized forest classifier. A pixel wise labeling is used to infer the 3D position of each joint.

The capability of the Kinect sensor to detect joints of the human body can be utilized to track a human being and follow him. But the robot should be able to detect and avoid any obstacles in its path during navigation. Peasley et al [18] proposes a novel approach for obstacle detection and avoidance.3D points are projected onto a ground plane and a 2D map is then used to detect the obstacles. This is then used to compute the translational and rotational velocities which are required to avoid the obstacle. An infinite pole approach and a control strategy are used to overcome the limitations of the sensor. The results obtained from experimenting in a variety of indoor environments display the robust obstacle avoidance in real time.

ARTIFICIAL NEURAL NETWORK FOR FACE RECOGNTINION

The Kinect can be used as the sensor for the robot navigation. Once a user comes into the field of view of the Kinect, it will be able to detect the presence of the user. In certain scenarios, it may be required to recognize who is the user and face recognition is suitable for identifying the user. Artificial neural networks are electronic models based on the neural structure of the brain, with neuron as the fundamental processing element. Dendrites, soma, axon and synapses are the four basic components of neurons. Dendrites act like input channels receiving input through the synapses of other neurons. The soma processes the incoming signals and turns the processed value into an output which is sent out through axon and synapses to other neurons. The artificial neural network simulates the behaviour of these four components. Pattern recognition can be accomplished very well with artificial neural network.

Neural networks are structured to provide the capability to solve problems without the benefits of an expert and without the need of programming. They can seek patterns in data that no one knows are there [19].

Face recognition from the images is challenging due to the wide variability of face appearances and the complexity of the image background. Nandini et al [20] proposes a facial recognition approach using Neural Network. It has Face localization part, where mouth end point and eyeballs will be obtained. In feature Extraction, Distance between eyeballs and mouth end point will be calculated. The recognition accuracy achieved by the method was very high.
Yang et al [21] uses image detection using ANN to classify the presence of weeds in cornfields. Images were taken from cornfields and given as input to the ANN. Many backpropagation ANN models were developed with different numbers of PEs in their hidden and various output layers. The performance of the ANNs was compared and the success rate for the identification of corn was observed to be as high as 80 to 100%, while the success rate for weed classification was as high as 60 to 80%.

CONCLUSION

For the robot to be used in different scenarios as an autonomous companion, the main aim is recognizing and tracking a human target using avoiding obstacles. The problems that have to be tackled are:

- Identifying a certain person and tracking only the identified person.
- The robot should be capable of understanding the commands of the user
- Once a user is identified, the robot should only follow the commands given by the primary user.
- The robot should be able to communicate with the user.
- Human target tracking while keeping a minimum distance to the target
- While following it should be able to detect any obstacle which may obstruct its path.
- Once an obstacle is detected, the robot should stop moving and then move again avoiding the obstacle.

For meeting these objectives, different mechanisms were explored. The Microsoft Kinect sensor is the best solution for this scenario. It helps to track the human and at the same time, the robot can follow the user avoiding obstacle.

REFERENCES:

1. AURÉLIE CLODIC , SARA FLEURY , RACHID ALAMI , RAJA CHATILA , GÉRARD BAILLY , LUDOVIC BRÈTHES , MAXIME COTTRET , PATRICK DANÈS , XAVIER DOLLAT , FRÉDÉRIC ELISEÏ , ISABELLE FERRANÉ , MATTHIEU HERRB , GUILLAUME INFANTES , CHRISTIAN LEMAIRE , FRÉDÉRIC LERASLE , JÉRÔME MANHES , PATRICK MARCOUL , PAULO MENEZES , UNIVERSITÉ PAUL SABATIER, "RACKHAM: AN INTERACTIVE ROBOT-GUIDE," IN ROBOT AND HUMAN INTERACTIVE COMMUNICATION, 2006. ROMAN 2006
2. R. Kirby, J. Forlizzi, and R. Simmons, "Natural person-following behavior for social robots," in *Proceedings of Human-Robot Interaction*, March 2007, pp.17–24.
3. S. Thrun, M. Beetz, M. Bennewitz, W. Burgard, A. Cremers, F. Dellaert, D. Fox, D. H¨ahnel, C. Rosenberg, N. Roy, J. Schulte, and D. Schulz, "Probabilistic algorithms and the interactive museum tour-guide robot Minerva," *International Journal of Robotics Research*, vol. 19, no. 11, pp. 972–999, 2000.
4. T. Breuer, G. Giorgana Macedo, R. Hartanto, N. Hochgeschwender, D. Holz, F. Hegger, Z. Jin, C. Mller, J. Paulus, M. Reckhaus, J. lvarez Ruiz, P. Plger, and G. Kraetzschmar, "Johnny: An autonomous service robot for domestic environments," *Journal of Intelligent and Robotic Systems*, vol. 66, no. 1-2, pp.245–272, 2012.
5. T. Wisspeintner, T. van der Zant, L. Iocchi, and S. Schiffer, "RoboCup@Home: Scientific competition and benchmarking for domestic service robots," *Interaction Studies*, vol. 10, no. 3, pp. 393–428, 2009.
6. A. Haasch, S. Hohenner, S. Huwel, M. Kleinehagenbrock, S. Lang, I. Toptsis, G. A. Fink, J. Fritsch, B. Wrede, and G. Sagerer, "Biron - the bielefeld robot companion," *in Proc. Int. Workshop on Advances in Service Robotics*, 2004, pp. 27–32.
7. M. Takahashi, T. Suzuki, H. Shitamoto, T. Moriguchi, and K. Yoshida, "Developing a mobile robot for transport applications in the hospital domain," *Robotics and Autonomous Systems*, vol. 58, no. 7, pp. 889–899, Jul. 2010.
8. U. Reiser, C. Connette, J. Fischer, J. Kubacki, A. Bubeck, F. Weisshardt, T. Jacobs, C. Parlitz, M. H¨agele, and A. Verl, "Care-o-bot R: creating a product vision for service robot applications by integrating design and technology," *in Proceedings of the IEEE Conference on Intelligent Robots and Systems (IROS)*,2009, pp. 1992–1998
9. N. Roy, G. Baltus, D. Fox, F. Gemperle, J. Goetz, T. Hirsch, D. Margaritis, M. Montemerlo, J. Pineau, J. Schulte, and S. Thrun, "Towards personal service robots for the elderly," in Proceedings of the Workshop on Interactive Robots and Entertainment (WIRE 2000), Pittsburgh, PA, May 2000.
10. H.-M. Gross, C. Schr¨oter, S. M¨uller, M. Volkhardt, E. Einhorn, A. Bley, C. Martin, T. Langner, and M. Merten, "Progress in developing a socially assistive mobile home robot companion for the elderly with mild cognitive impairment," in Proceedings of the IEEE Conference on Intelligent Robots and Systems (IROS), 2011, pp. 2430–2437
11. Aggarwal, J.K.andCai, Q. "Human motion analysis: a review" in Nonrigid and Articulated Motion Workshop,San Juan , June 1997.
12. Donato Di Paola, Annalisa Milella, Grazia Cicirelli and Arcangelo Distante" An Autonomous Mobile Robotic System for Surveillance of Indoor Environments *" International Journal of Advanced Robotic Systems*, Vol. 7, No. 1 (2010)
13. Maria C. Garcia-Alegre, David Martin, D. Miguel Guinea and Domingo Guinea" Real-Time Fusion of Visual Images and Laser Data Images for Safe Navigation in Outdoor Environments" in *Sensor Fusion - Foundation and Applications,Published*: June 13, 2011
14. Z. Chen and S. T. Birchfield, "Person following with a mobile robot using binocular feature-based tracking," in *Proceedings of the IEEE Conference on Intelligent Robots and Systems (IROS)*, Oct. 2007.

15. Luca Marchetti , Daniele Pucci, Pascal Morin "Autonomous Shopping Cart Platform for People with Mobility Impairments" in *Workshop at IROS'12 on Assistance and Service robotics in a human environment*., Oct 2012, Villamoura, Portugal.

16. K. Khoshelham "Accuracy analysis of kinect depth data" in *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume XXXVIII-5/W12, 29-31 August 2011, Calgary, Canada

17. J. Shotton, A. W. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman, and A. Blake, "Real-time human pose recognition in parts from single depth images," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011, pp. 1297–1304.

18. Peasley, B and Birchfield, S., "Real-Time Obstacle Detection and Avoidance in the Presence of Specular Surfaces Using an Active 3D Sensor," in Robot Vision (WORV), IEEE Workshop, 15-17 Jan. 2013, Clearwater Beach, FL

19. http://www.psych.utoronto.ca/users/reingold/courses/ai/cache/neural2.html

20. M.Nandini, P.Bhargavi, G.Raja Sekhar, "Face Recognition Using Neural Networks," in *International Journal of Scientific and Research Publications*, Volume 3, Issue 3, March 2013

21. C.-C. Yang, S.O. Prasher, J.-A. Landry, H.S. Ramaswamy and A. Ditommaso, "Application of artificial neural networks in image recognition and classification of crop and weeds" *NY, USA, June 2000*