Target Detection and Tracking using Intelligent Wheelchair

Saima Rehman [1], Tabinda Ashraf [2], Muhammad Umair [3], Usama Zubair [4], Yasar Ayaz [5], Hammad Khan [6]

Robotics and Intelligent Systems Engineering (RISE) Lab
Department of Robotics and Artificial Intelligence
School of Mechanical and Manufacturing Engineering (SMME), & College of Electrical and Mechanical Engineering,
National University of Sciences And Technology (NUST), H-12 Campus
Islamabad, 44000, Pakistan.
[1] saimarehman12@gmail.com; [2] tabi_ash74@hotmail.com; [3] umair6428@gmail.com; [4] osamazbr232@gmail.com;
[5] dr.yasar.ayaz@gmail.com; [6] khammad88@yahoo.com

Abstract **- Human-robot interaction is one of the most basic requirements for service robots. In order to provide the desired service, these robots are required to detect and track people in human cluttered environment. Once the target to be followed has been recognized by robot, next step is navigation while taking care of obstacles. This paper presents a novel approach for the target detection to make the robot able of target tracking in a cluttered environment. Using data from Hough transform our system classifies the target person from other human beings in the environment. Our system tracks human being by gathering details of his position and velocity, and then converting this data into corresponding linear and angular velocity of robot. System used in the project is an intelligent wheelchair with partial AI and partial human command based working architecture with LRF and stereo camera mounted on it.**

Keywords - Target detection and tracking; Circle Hough Transformation; Obstacle Avoidance; Motion Planning

I. INTRODUCTION

The development of service robots that are able to assist people in their daily tasks has become a popular research area in recent years. Possible application scenarios for such robots include support for the elderly and the general service robots for public areas or shopping malls. An essential prerequisite for service robots is the ability to recognize humans and interact with them in a non-technical, natural way.

The primary objective of this project is to promote the development of robust, repeatable and transferable software for robots that automatically detect, track and follow people around them in real life with number of hurdles present in work area. The work will be developed by the necessity of such functionality onboard on intelligent electric wheelchair robots to assist people with mobility impairments. Our approach is effective for different robot platforms. The method has been implemented in the Robot Operating System (ROS) framework and will be publicly released as ROS package. The research is carried out in collaboration with Sakura Wheelchair project (Japan).

The intelligent wheelchair project focuses on improving the lives of people with severe disabilities who are not in the situation of directing electric wheelchairs on their own. The aim of the project is to make wheelchair users able to autonomously navigate using advanced environment perception, motion planning and navigation techniques through everyday obstacle cluttered human environments.

The ability to autonomously detect, track and follow a person has been identified as an important functionality for many assistive and service-robot systems. In the last decade, significant advances have been made in the development of people detection and tracking algorithms, which are often carried out with the aim of improving the human-robot interaction and robot navigation in populated environments.

But most of the previous work is not easily transferable to new applications: algorithms are tested on a single robot in a single environment (if at all, sometimes only in the simulation under artificial conditions), in many cases, the code has not been made available for public, records collected in the validation sessions are not shared and quantitative comparisons to existing algorithms has not been performed.

In this project we focus on detection of a target and tracking it as important, fundamental interaction capability for a service robot equipped with a color camera and a laser scanner, while wheelchair is following its care taker avoiding the obstacles in the way.

II. LITERATURE SURREY AND PREVIOUS WORK

Generally speaking, a wheelchair is capable of sensing its caretaker, location, interpreting the sensed information to obtain the knowledge of its location and direction of motion, planning a real-time trajectory to reach the object. In this process, the issue of recognition of target is a fundamental topic to be challenged. It is one of the most fundamental problems that have to be solved before the mobile robot can navigate and explore autonomously in complex indoor environment. The requisite target once detected make way to achieve the task of generating a feasible path which can guarantee mobile robot's moving from start to the target safely and optimally.

In the wheelchair project, we have utilized data from the camera mounted on wheelchair to identify marker of the target and move along a path to follow it.

Different areas of research, such as parking, navigation, trajectory tracking, and wall/lane-following, are all actively being investigated. Nowadays, mobile robots are a promising technology that will improve the quality of life by providing collision avoidance, reducing traffic gridlock, and allowing the replacement of dangerous tasks currently performed by human drivers. The path planning for the objective of obstacle avoidance which is one of the key issues in the mobile robot navigation, is the major topic to be investigated in this study. SO the first addressed issue is the fair and error free detection of the target. Though a lot of work has been done in this regard, various algorithms have been designed but the need of this project is to detect the care taker efficiently and then follow him in his footsteps by planning the path and avoiding the obstacles. Most previous work focuses on only one or two of the three identified sub-problems: detection, tracking and following (with obstacle avoidance and path planning). Few papers present integrated systems tackling all three components. We focus in particular on work that uses depth sensor, e.g. laser range finder, and visual camera, as these are the sensor available on our Smart Wheeler robot, as well as on numerous other assistive robots.

A. Stereo Vision

Jia et al. [1] presented a human detection and tracking approach using stereo vision and extended Kalman filter-based methods. Kristou et al. [2] presented their approach for identifying and following the targeted person using a laser rangefinder and an omnidirectional camera. Their method fused the data of both these sensors by identifying the target using a panoramic image, and then tracking that targeted person using the laser rangefinder. Recently, Petrović et al. [3] presented a real time vision-based tracking method using a modified Kalman filter. They used a stereo vision-based detection method to get the features from 2D stereo images, and then reconstructed them into 3D object features to detect human beings. These techniques are suitable when only one person is present in the environment the target person and it is not able to handle more than one person in the environment. Schulz et al. [4] proposed to estimate the number of people in the current scan based on the number of moving local minima in the scan. Unfortunately, this requires people move continuously to be tracked, and is susceptible to poor results in cluttered environments. Topp et al. [5] extended [4], by picking out shapes of legs and person-wide blobs in laser scans using hand-coded heuristics, to allow detection and tracking of both stationary and moving people. The approach was also combined with a person following navigation algorithm, combining both the tracked person's position as well as the location of nearby obstacles to determine suitable control. Gockley et al. [6] used a similar approach, with a few modifications, including using a Brownian motion model for the tracking component. This approach was further extended by Hemachandra [7], which improved the person-following component by proposing a navigation approach that accounts for personal space, while avoiding obstacles. Unfortunately these approaches cite tracking difficulties in cluttered conditions, since they relied primarily on detecting clusters of a heuristically determined size in the laser scan.

B. Supervised learning

More recently, Arras et al. [8] reduced this limitation by proposing a method that detects legs by first clustering scan points and then using supervised learning to learn shapes of leg clusters. Detected legs are tracked over time using constant-velocity Kalman filters and a multiple-hypothesis tracking (MHT) data association technique. This approach benefits from its ability to maintain (but not initiate) tracks of stationary people, and does not require any priori occupancy grid map of the environment. Initial results for this method appear promising, but demonstrations on walking speed robots in cluttered and crowded areas have yet to be performed. Thus, many questions remain about the robustness and generalizability of the approach.

C. Omni-Directional Camera

Kobilarov et al. [9] developed a person-following Segway robot using an Omni-directional camera and a laser scanner. Nonetheless, variable environment lighting, backgrounds and appearances of people are factors which can be difficult to control for and can mislead vision-reliant tracking systems.

Fast and accurate detection of circles is widely applied in the fields of image processing and computer vision. The Hough transform (HT), as a basic method to detect circles, has its advantages in insensitivity to noise in images and easiness in parallel computing, which has attracted extensive research among researchers. HT has the advantages of small storage and high speed if the parameter space is limited to two dimensions. Based on this we decided to detect the caretaker based on a circular tag present of his/her shirt rather than detecting the person himself.

D. Circular Shapes Detection

Sambarta Dasgupta [10] presented an algorithm for the automatic detection of circular shapes from complicated and noisy images without using the conventional Hough transform methods. The proposed algorithm is based on a recently developed swarm intelligence technique, known as the bacterial foraging optimization (BFO). A new objective function has been derived to measure the resemblance of a candidate circle with an actual circle on the edge map of a given image based on the difference of their center locations and radii lengths. Guided by the values of this objective function (smaller means better), a set of encoded candidate circles are evolved using the BFO algorithm so that they can fit to the actual circles on the edge map of the image. The proposed method is able to detect single or multiple circles from a digital image through one shot of optimization.

T. D`Orazio [11] introduced a circle detection algorithm based on the CHT that has been formulated as convolutions applied on the edge magnitude image. The convolution kernels have been properly defined in order to detect the most completed circle in the image, being independent on the edge magnitude, and considering also different shapes of the ball according to different light conditions.

In 2005, Victor Ayala-Ramirez [12] presented a circle detection method based on genetic algorithms. It reduces the search space by avoiding trying unfeasible individuals, this results in a fast circle detector. The approach detects circles with sub-pixellic accuracy on synthetic images, but it can also detect circles on natural images. It also handles detection of occluded circles and several circles in single image.

T. J. Atherton [13] showed that a specific combination of modification to the CHT is formally equivalent to applying a scale invariant kernel operator.

Here is an overview of working of our system:

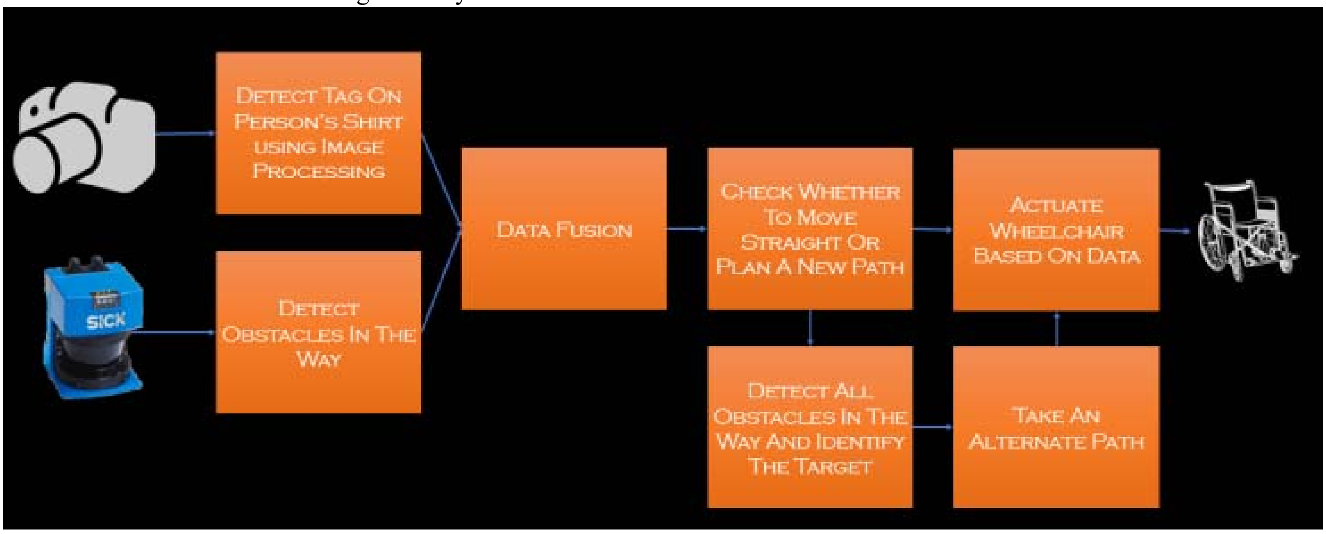

Figure 1: Block Diagram of System

III. HUMAN DETECTION

The first stage of each visual system is the image acquisition stage. The image being captured is completely unprocessed and is the result of what hardware was used to generate it. After the image has been obtained, various methods of processing can be applied to the image to perform many different visual tasks that are required. Here are few of the processing techniques that we have applied:

- RGB to HSV conversion
- Image Smoothing
- Thresholding
- Hough Circle Transformation

In order to detect a human, we had various options available:

- Face detection
- Human detection
- Upper body detection
- And many more

But we have not chosen any of them since these detection/ recognition techniques are computer-intensive and the processing time becomes a problem. This time taken by this processing is so much that the detection technique cannot be applied to frames of a video stream in real time. In order to save time, we have decided to recognize a particular shape on the shirt of the caretaker, for simplicity the label is assumed to be a smooth circle with a red color.

A. RGB to HSV Conversion

Images can be processed either in the RGB format (red, blue and green) or HSV (Hue, Saturation and Value). HSV is used when the color description has an important role to play. It describes colors in a similar way as a human eye tends to recognize colors. It uses the well-known comparisons for description such as color, vibrancy and brightness.

- The hue represents the color type whose value is normalized to a range of 0 to 255, where 0 is red.
- Saturation represents the liveliness of the color. Its value ranges from 0 to 255. The lower the saturation value, the more gray is present in the color so that it appears faded.

- Value represents the brightness of the color. It ranges from 0 to 255, where 0 is completely dark and 255 is full.
- White has an HSV value of 0-255, 0-255, 255. Black has an HSV value of 0-255, 0-255, 0. The dominant description for black and white is the term value. The hue and the saturation value do not make any difference when the value is at the maximum or minimum intensity value.

The color camera uses the RGB model to determine the color. Once the camera has read these values, they are converted to HSV values. The HSV values are then used in the code to determine the specific color. The pixels are individually examined to determine if they match a predetermined color threshold.

B. Image Smoothing

A simple and frequently used OpenCv operation is the image smoothing. It is also referred to as blurring the image. We need to blur or smooth the image to reduce the noise.

To smoothing an image, we use a filter; usually, filters are used in which the weighted sum of the input pixels f(i, j) determines the output pixels g(i, j) [1-2].

$$g(i,j) = \sum_{k,l} f(i+k, j+l) * h(k,l) \quad (1)$$

The second term h(k,l) in the summation is the kernel, i.e., the coefficients of the filter. This is useful to visualize the filter as a sliding window of coefficients over the image.

The most commonly used filter we have implemented is the Gaussian filter. The filter operates by folding each point in the input array to the Gaussian kernel, followed by the addition of all to give the output array.

C. Thresholding

Since the noise content in the image has been suppressed, it can be further processed. The first processing we do is threshold. Since color value is obtained due to reflection and brightness, is not always the same. So we define a range of HSV values for our tag, and threshold the image:

$$f(x,y) = \begin{cases} 1 & if\ HSV(x,y) \in HSV\_tag \\ 0 & otherwise \end{cases} \quad (2)$$

Where a 1 is the white color, while 0 is the black color. HSV (x, y) is the HSV value of the pixel to be processed, while HSV_tag is the defined range of HSV values by day [2]. Threshold image can be seen in figure 5.

D. Circular Hough Transform

Among the most sophisticated techniques of image processing and computer vision, determining the shape, location or amount of appearance is the one of the focus and concerns in many applications. In our approach to human detection using a circular marker, we decided to use the most robust technique of Circular Hough Transform (CHT). The disagreements with the change of the viewing angle, the range and the lighting environment are indicated by the information of the circle target.

The Hough transform-based standard method for finding circles has the advantages of a strong noise resistance, a strong robustness against the interference point, a low sensitivity to edge defects, and an excellent adaptability for parallel algorithms. The problems of the enormous calculation and the large time-consumption in the 3D projections of the boundary image in the parameter space such as useless accumulation, quantization errors were solved by implementing the Circular Hough Transform Technique. It was ensured that the demands on timeliness and truthfulness, if the background is too complex, are fulfilled. The Hough transform is used to determine the parameters of a circle because a number of known points fall on the parameter. The parameter space is 3D; if the circles in an image have the known radius R, the search can be reduced to 2D. The circle with Radius R and a center (a, b) can be expressed using the parametric equations as follows:

$$x = a + R\cos\theta \quad (3)$$

$$y = b + R\sin\theta \quad (4)$$

The location of (a, b) -points in the parameter space falls on a circle with the radius R centered on (x, y). The true center is common to all parameter circles and can be found with a Hough accumulation array. While the angle arcs through the entire 360 ° range, the coordinates (a, b) track the circumference of the circle. A search program is then used to find the triple parameters to facilitate the description of the circle, e.g. R, a, b in the image containing many dots, some of which fall on the circumference.

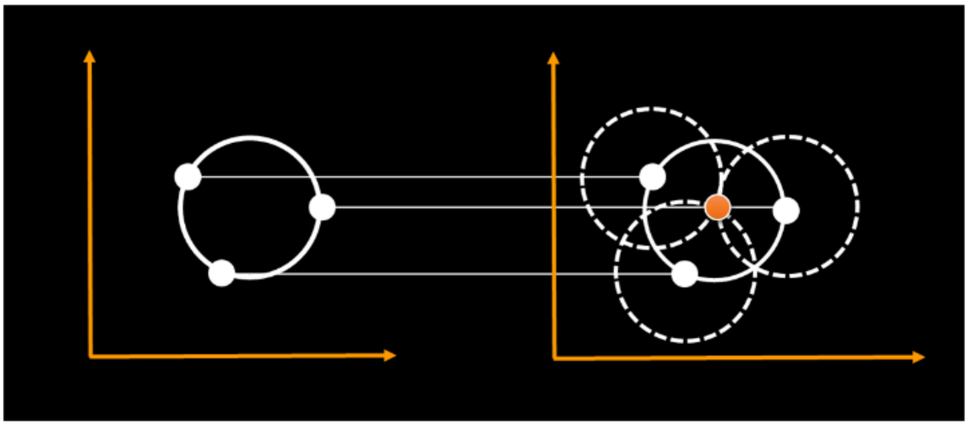

Figure 2: Understanding basic theory of circle Hough transform

A three dimensional array is used by the circle Hough transform technique where the first two elements represent the coordinates of the center of circle while the third is the value of radius. The values in the array keep increasing every time a circle is drawn. The accumulator keeps counting the number of circles passing through the coordinates of each edge point this gets proceeded to a vote to find the highest count and the circles seen in images are those coordinates which have highest count.

In our project R is not predefined or fixed, it varies with forward and backward movement of the caretaker. Since the caretaker is usually in the 2-3m range of the wheelchair, we have measured the radius for the nearest and the farthest point. The fact that the parameter space is 3D makes a direct implementation of the Hough technique, which is more expensive in terms of computer memory and time. Limit of the radius increased speed. The technique was used for live video from the camera.

Several circles of equal radius can be found with the same technique. The center points are displayed as red cells in the parameter space drawing. An overlapping of circles can also cause defective centers, such as the blue cell. Spurious circles can be removed by matching circles in the original image.

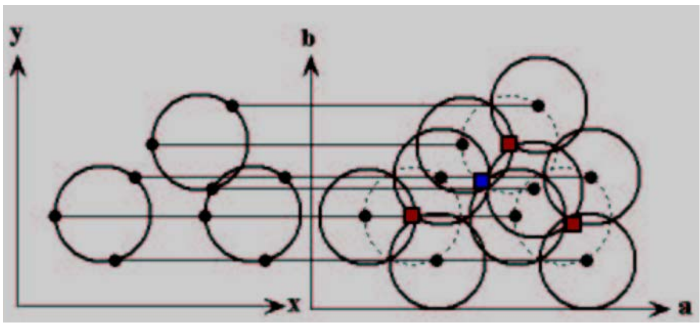

Figure 3: Defective centers shown in blue

Since the radius is not fixed, the location of the points in the parameter space will fall on the surface of a cone. Each point (x, y) at the circumference of a circle creates a conical surface in the parameter space. The triplet (a, b, R) corresponds to the accumulation cell, in which most of the cones intersect. The following figure shows the creation of a conical surface in the parameter space for a (x, y) point. A circle with a different radius is built up on each plane. The search for circles with an unknown radius can be carried out using a three-dimensional accumulation matrix.

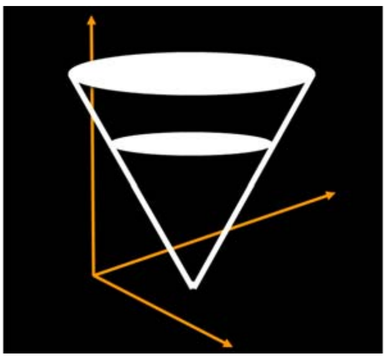

Figure 4: 3D Parameter space

E. Amendments in Transform

Applying whole algorithm on each and every frame of the video was time consuming. Also the variation in the results was occurring so fast that it became impossible for wheelchair to properly track the target. So we applied Hough transform on multiple frames, took the mean of the radius and center of the recognized circle and used these mean values to command wheelchair.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

This section focuses on experimental results of our algorithm. We have tested this project for multiple scenarios. In all of them, circle was detected to accuracy required, as shown in diagram below.

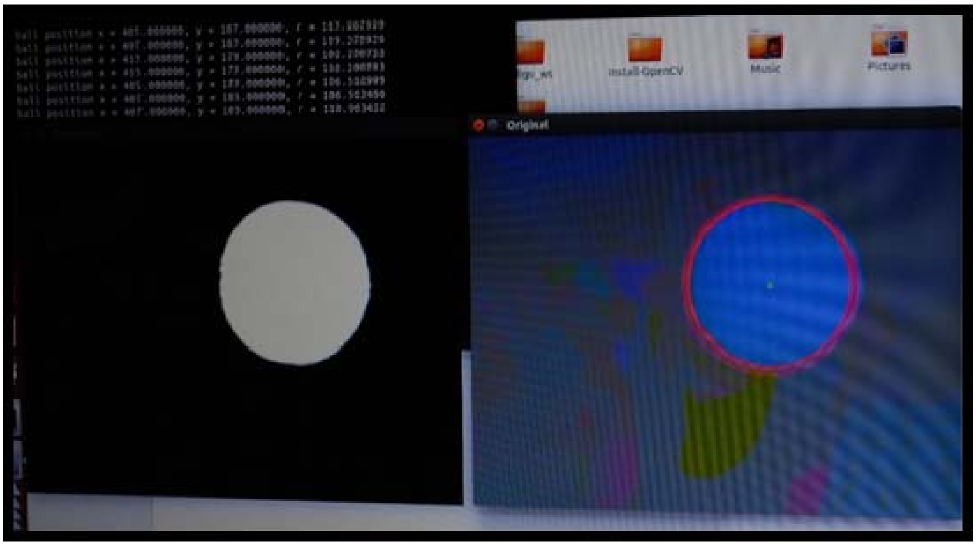

Figure 5: Result of Circle detection. On left hand side is the threshold-ed image, while right hand side shows result with circle detected and highlighted in red color

A target was detected in two different scenarios:

1. Indoor

Wheelchair perfectly tracks caretaker in indoor environment.

2. Outdoor

In the outdoor environment, performance depends on light intensity incident on the marker.

Case 1:

Initially, we considered a maze type environment, and tracked the path followed by target and wheelchair. As seen from result, the wheelchair has followed the caretaker quite efficiently.

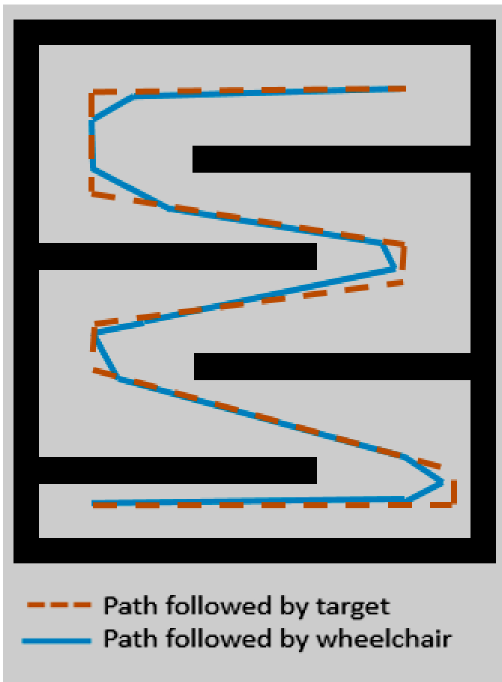

Figure 6: Result for case 1

DOI 10.5013/IJSSST.a.18.01.04 ISSN: 1473-804x online, 1473-8031 print

Case 2:

Here we considered static obstacles that target can jump over but wheelchair can't. So it plans alternate path to follow the target.

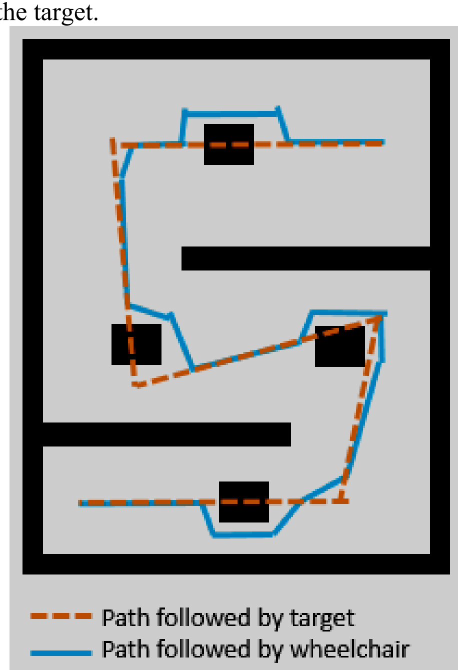

Figure 7: Result for case 2

Case 3:

Now if obstacle is dynamic, the wheelchair follows a path from behind the obstacle, so as to facilitate the motion of both the obstacle and wheelchair.

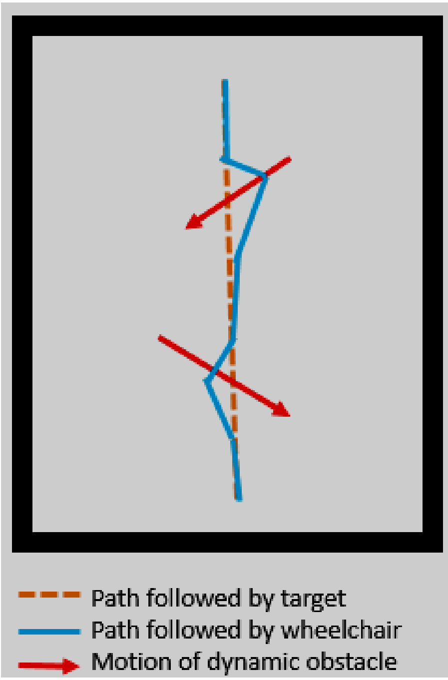

Figure 8: Result for case 3

V. CONCLUSION AND FUTURE WORK

The paper focused on the efficient techniques of human detection and following. It detects a person based on the tag printed on its shirt. Based on the features collected via camera, it measures the distance and location of target human; and generates corresponding velocity commands to the motors of the wheelchair. A series of algorithms were combined and implemented on wheelchair of RISE Lab, SMME; results proved that it can be used on commercial scale. Our focus was on static obstacles and just one dynamic obstacle. The tag on the shirt of caretaker can be made to include some text i.e. serial number of the chair, so that wheelchair crowded environment every chair can recognize its caretaker. The task done now, uses webcam of laptop, instead of this, already mounted camera, or a new camera can be added to it, for permanent and dedicated use.

ACKNOWLEDGMENT

The authors are grateful to Dr. Hong Zhang for sharing their work on Person Tracking and Following with 2D Laser Scanners. Moreover, authors are also thankful Badar Ali for insightful discussion marker based target detection using digital image processing.

REFERENCES

[1] Jia, S., Zhao, L., & Li, X. 2012. Robustness improvement of human detecting and tracking for mobile robot. In Mechatronics and Automation (ICMA), 2012 International Conference .pp. 1904 1909. IEEE.

[2] Kristou, M., Ohya, A. & Yuta, S. 2011. Target person identification and following based on omnidirectional camera and LRF data fusion, 20th IEEE International Symposium on Robot and Human Interactive Communication. Pp.419-424.

[3] Petrović, E., Leu, A., Ristić-Durrant, D. & Nikolić, V. 2013. Stereo vision based human tracking for robotic follower International Journal of Advanced Robotic Systems. 10:230 DOI: 10.5772/56124

[4] D. Schulz, W. Burgard, D. Fox, and A. B. Cremers, "People tracking with mobile robots using sample-based joint probabilistic data association filters,"Int. Journal of Robotics Research, 2003.

[5] E. A. Topp and H. I. Christensen, "Tracking for following and passing persons" in Int. Conference on Intelligent Robots and Systems (IROS), 2005.

[6] R. Gockley, J. Forlizzi, and R. Simmons, "Natural person-following behavior for social robots," in Int. Conference on Human-robot Interaction, 2007.

[7] S. Hemachandra, T. Kollar, N. Roy, and S. Teller, "Following and interpreting narrated guided tours," in Int. Conference on Robotics and Automation (ICRA), 2011.

[8] K. O. Arras, B. Lau, S. Grzonka, M. Luber, O. M. Mozos, D. MeyerDelius, and W. Burgard, "Range-based people detection and tracking for socially enabled service robots," in Towards Service Robots for Everyday Environments, pp. 235–280, 2012.

[9] M. Kobilarov, G. Sukhatme, J. Hyams, and P. Batavia, "People tracking and following with mobile robot using an omnidirectional camera and a laser," in Int. Conference on Robotics and Automation (ICRA), 2006.

[10] Sambarta Dasgupta & Swagatam Das, 'Automatic circle detection on digital images with an adaptive bacterial foraging algorithm'/ October 15, 2009

[11] T. D'Orazio,C. Guaragnella,M. Leo,A. Distante,' A new algorithm for ball recognition using circle Hough transform and neural classifier, Volume 37,Issue 3,March 2004

[12] Victor Ayala-Ramirez & Carlos A Gracia,'Circle detection on images using genetic algorithms, University of Gwantawa , Electronics and communication engineering department, Mexico/November 28,2005

[13] T J Atherton & D J Kerbyson. 'Size Invariant Circle Detection', Department Of Computer Science, University Of Warwick