

# Depth based Tracking of Moving Objects using Kinect

\*Mahanthesha U ,Pooja ,<sup>2</sup>Rakshitha R ,Sherly Vincent , Smitha S Potdar

Department of Electronics& Communication Engineering

GSSS Institute of Engineering & Technology For Women, Mysuru

**Abstract**— This project presents a fire bird V robot with a Kinect sensor as its only sensor uses image processing to track the target and avoid collision. It is the goal of the study to eliminate certain limitations of robots such as limited operating area and requiring the user to wear intrusive sensors. The robot detects a person based on depth estimation. Several tests are made to characterize the robot, these test are: straight path test, varying curve radius test, and collision avoidance test. Though the robot was able to follow the user in a straight path, it was noticed that there is an average deviation of 4 to 8 cm. from the path of the user. In the varying curve test, the safest turning radius is at least 1.5 meters, which enables the robot to not collide with the vertex of the arc made by the user and not lose track of the user. In this project, the robot also tends to undercut the user. For collision avoidance tests conducted, the robot was successful to avoid obstacles but it makes, at most, a 180 cm. of deviation from the path of the user. Although the robot is able to follow a user in a straight path; the robot tends to undercut the user during curved path.

**Key words**— Kinect sensor, FIRE BIRD V Robot

## I. INTRODUCTION

Robotic technology has increased appreciably in past couple of years. Such innovations were only a dream for some people a couple of years back. But in this rapid moving world, now there is a need of robot such as “A Human Following Robot” that can interact and co-exist with them. To perform this task accurately, robot needs a mechanism that enables it to visualize the person and act accordingly [2].

The image processing carried out to get the information about the surroundings visually is a very important thing. The following points should be carefully noted while doing the processing.

Typically human following robots are equipped with several different diverse combination of sensors i.e. light detection and ranging sensor, radio frequency identification module (RFID), laser ranger finder (LRF), infrared (IR) sensing modules, thermal imaging sensors, camera, wireless transmitter/receiver etc. for recognition and locating the target. All the sensors and modules work in unison to detect and follow the target. In this project we use Kinect sensor which include above mentioned criteria and fire bird V as the robot. In this project, we presented a method of human following robot based on depth estimation and detection by using two inbuilt cameras. Intelligent tracking of specified target is carried out by the

use of different sensors and modules [5] i.e. ultrasonic sensor, magnetometer, infrared

sensors and camera. An intelligent decision is being made by the robot control unit based on the information obtained from

the above sensors and modules, hence finding and tracking the particular object by avoiding the obstacles and without collision with the target.

The paper is organized as follows: Section II gives details of hardware components of the robot system and how these components are implemented. Section III focuses on software development. It gives details on how the Kinect tracks users and robot follows accordingly. In Section IV, it suggests the future possible directions and improvement. Lastly, Section V leads to conclusion.

## II. OBJECTIVES

- To build a smart robot that can automatically follow a person around in a room.
- Robot detects the position of person and calculate the distance using IR sensors and cameras.
- To implement a buzzer when the target is lost.

## III . LITERATURE SURVEY

1. REVIEW ON MICROSOFT KINECT’S GESTURE TECHNOLOGY BASED ROBOTICS- Praveen Manoharan, Nithin Sundararaj, Udhaya Sankar and S J Syed Ali Fathima in 2017

The field of sensor based human computer interaction and robotics is rapidly developing. A Microsoft Kinect device to implement a suitable interface between a machine and human where users can control machines (for e.g. Computer, robot etc.). The skeletal tracking ability of Kinect provides an efficient way of controlling a device through gestures. The Kinect sensor uses a depth map and colour image as well. Gesture based human computer interaction is an emerging domain of research. This paper introduces a basic understanding on the gesture based technologies and how we can implement a robotic arm using a device known as Kinect. The skeletal tracking ability of Kinect plays a vital role in this.

## 2. SCENE PARAMETERS ANALYSIS OF SKELETON BASED HUMAN DETECTION FOR MOBILE ROBOT USING KINECT- S.A. Abdul Shukor\*, Mohammad Amir Abdul Rahim, B. Ilias in 2016

In this paper, the Kinect is being used as a tool for human detection, which is an essential task before tracking. This paper describes the best parameters that can be utilized in human tracking using skeleton-based method. In this paper, different types of parameters are quantitatively measured in order to choose the optimum scene that can be used for tracking human using skeleton-based method. Parameters such as distances to the sensor, lighting intensity, and the presence of obstacle are selected as these are among the main factors that need to be considered while tracking human indoor. The ideal conditions for the Kinect to detect and track human using skeleton-based algorithm would be at a distance of 3 meters from the sensor, lightings of 158 lux, and with minimal or no obstacles at all.

## 3. TELEOPERATION CONTROL OF BAXTER ROBOT BASED ON HUMAN MOTION CAPTURE- Guangzhu Peng, Chenguang Yang\*, Yiming Jiang, Long Cheng and Peidong Liang in 2016

Currently, vision based motion capture systems provide us an alternative approach for motion capture technology. In this paper, based on human kinematics and function approximation technique (FAT), a novel method is presented for the trajectory control of Baxter robot. Each arm of the Baxter robot has seven degrees of freedom. The geometry vector approach is applied to capture human motion trajectory by using Microsoft Kinect sensor. A FAT control system is employed to make the robot follow the trajectory of human motion. The UDP communication protocol is employed to send the reference human joint angle data to the robot.

## 4 A NOVEL APPROACH TO A MOBILE ROBOT VIA MULTIPLE HUMAN BODY POSTURES – Dajun Zhou, Fei Chao, Zuyuan Zhu, Chih-Min Lin, and Changle Zhou in 2016

This paper focuses on applying human postures and face tracking technologies to design an autonomous patrol vehicle control system, which contains a wireless video surveillance ability. The system includes following parts: (1) Obtaining the skeleton joints based on the Kinect skeleton tracking: the angles and distances between a human arm's joint are calculated to be the input of an artificial neural networks. (2) Changing the vehicle's gear box: the dynamic gesturers cognition is built by using the artificial neural network and a finite state machine. (3) Controlling the vehicle speed, the speed is controlled by a fuzzy control algorithm. (4) Controlling of a motorized camera: the Kinect face tracking function is applied to detect a human's face direction, so that the motorized camera's rotation is controlled by the direction.

## 5 REAL- TIME TARGET TARGETING SYSTEM FOR PERSON FOLLOWING ROBOT – Qimin Ren1, Qingjie Zhao1 ,Hui Qi1, Lingrui in 2016

In this paper, A very fast and robust tracking system for the person-following robot. The robot tracking system could detect human automatically in the field of view. The user issues a command by hand gesture as a flag of start, then the person-following robot locks the user as the target and starts tracking. The robot system consists of two parts: a basic tracker which uses the skeletal tracking algorithm that is provided by Kinect SDK and an auxiliary tracker which utilizes Cam shift algorithm. When the basic tracker fails, the auxiliary tracker utilizes Cam shift to correct the wrong result to ensure the robot obtains the right location. After getting the location of the target, we predict the position of next moment by the Extended Kalman filter.

## 6 EFFICIENT PEOPLE SEARCH AND GESTURE RECOGNITION USING A KINECT INTERFACED MOBILE ROBOT – Robert Tubman and Khalid Mahmood Arif in 2016

This paper describes the programming and application of human searching and gesture recognition using a neck-like swivel mechanism attached to a mobile robot Pioneer 3-DX interfaced with a Microsoft Kinect sensor. We utilize color, depth and skeleton data streams from the Kinect to find and track people for gesture commands to do basic robotic movements like wander around, follow-me etc.

## IV. METHODOLOGY

The robot control system is mainly comprised of a mobile robot platform, a Kinect camera, a computer (laptop), Zigbee module for communication. The system structure is shown in Fig. 1.

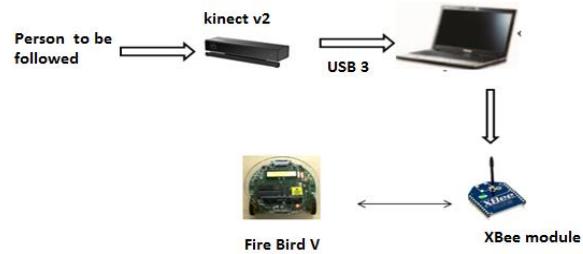


Fig.1 Block diagram

In the system, the computer acquired depth information using the Kinect camera, to achieve the functions of tracking the human body. Depth data stream is obtained and processed by the computer. And then it is transmitted to the sensor drive board through the serial bus. Finally the control commands are sent to control the robot to track people. The obstacle avoidance sensors can protect the mobile robot platform from damages caused by the collision between the robot and the object.

## KINECT

Kinect is a line of motion sensing into devices that was produced by Microsoft ,which consists of an RGB camera ,an infrared emitter, an IR depth sensor ,a multi array microphone and a tilt motor.



Fig2:kinect sensor

Depth camera: Analyzes IR patterns to build a 3-D map of the room and all objects and people within it.

Colour camera: like a webcam this captures a video image kinect uses that information to get details about objects and people in the room.

IR Emitter: projects a pattern of infrared light into the room. As the light hit a surface ,the pattern becomes distorted ,and the distortion is read by the depth camera.

Microphone array: Four mics pinpoint were voices or sounds are coming from while filtering the background noise.

Tilt motor: Automatically adjusts based the object in front of it.if you are tall it tilts the box up, if you are short it knows to angle down .

### Controller and Converter – Brain

A portable laptop of Windows system is used as the controller of the system. In order to transmit commands from the laptop to the robot system, and to read feedback information from the robot system, Zig bee module is used. It is able to generate PWM signal, read/ write input and output signals.

## FIRE BIRD V



Fig.3: Fire Bird V

Fire Bird V is a Universal Robotic Research Platform, provides an excellent environment for experimentation, algorithm development and testing. Fire Bird V is evolved

from Fire Bird IV and Fire Bird II which are being used in IIT Bombay to teach embedded systems and robotics. Its modular architecture allows you to control it using multiple

Joint 1	Hip center	Joint 2	Spine
Joint 3	Shoulder center	Joint 4	Head
Joint 5	Shoulder left	Joint 6	Elbow left
Joint 7	Wrist left	Joint 8	Hand left
Joint 9	Shoulder right	Joint 10	Elbow right
Joint 11	Wrist right	Joint 12	Hand right
Joint 13	Hip left	Joint 14	Knee left
Joint 15	Ankle left	Joint 16	Foot left
Joint 17	Hip right	Joint 18	Knee right
Joint 19	Ankle right	Joint 20	Foot right

processors such as 8051, AVR, PIC and ARM7 etc. Modular sensor pods can be mounted on the platform as dictated by intended applications. Precision position encoders makes it possible to have accurate position control. The platform can be upgraded to tank drive and Hexapod insect or any other desired form very easily. It is powered by high performance rechargeable NiMH batteries. A 2.4 GHz ZigBee module provides state of the art secure and multi-channel wireless communication up to a range of one kilometer.

In fire bird v six important module are present such as power management module, Sensing module, Modules for locomotion, Peripheral modules ,communication module and intelligence module. Intelligence modules are basically microcontrollers. ATMEGA 2560 and ATMEGA8 are present ,prior acts as master and other acts as a slave.

## FLOW CHART

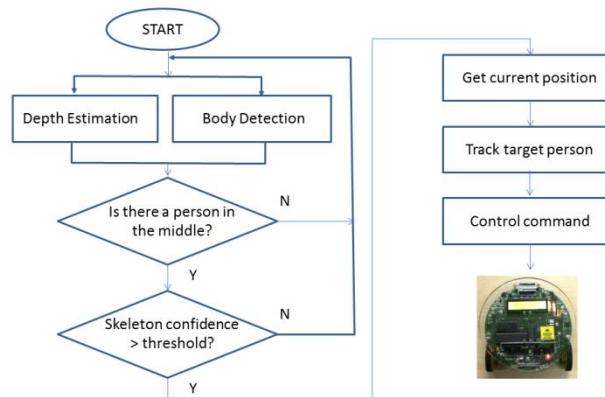


Fig.4: Flow chart

## BODY DETECTION

when the program finds a targeted person, it is not sufficient to tell if it is a human. According to experiments, false detections like mistreating a door, wall or a chair as a human. In order to detect valid human, skeleton detection confidence is introduced.

In Figure 5, in case like target person is too near, and only part of the body is detected, skeleton pattern is very

disordered. In this situation, it is not confident to tell the exact position of the user. Then program should not execute.

In order to avoid this problem and improve the accuracy, minimum 20 skeletal joints should be detected. They are listed in the table1.



Fig.5:skeleton detection

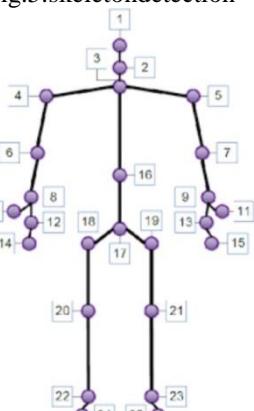


fig.6: human skeleton joint Representation

For each joint, there are 3 statuses to illustrate the confidence level. Tracked – it is confident to tell the joint is found. Inferred – it is not confident to tell the joint is found (may or may not exist). Not tracked – the joint is not found. Count the number of tracked joints to represent the confident level. The higher number, the more confident it is. There is another condition that the algorithm can be further improved. In this project, as when the user is close to the Kinect, joints of lower body is usually unable to be captured, which causes the confident level low. But in this case, the user does exist and program should execute. Thus count the number of upper body joints will be more accurate. Joints 14- 20 of the whole body will be omitted.

#### DEPTH ESTIMATION

The IR emits infra-red lights into the 3D space, and the light passes through the diffusion sheet to distribute within the measurement space. When the infra-red lights project on the body, the reflective spots are formed. The Projector records these reflection spots, and then synthesizes a depth

image with depth information by the chip. Access the original data streams from the depth sensor (IR & Projector), and obtain the depth information processed by the computer.



fig.7: color representation for various depth intensities

#### ANGLE CALCULATION

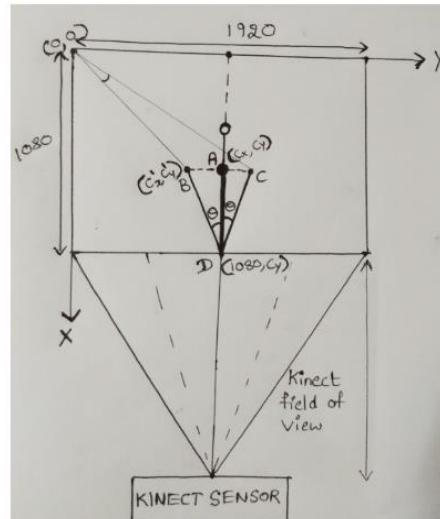


Fig.8: Angle calculation

Consider the point A, as the centre of field of Kinect where the angle is taken as  $0^\circ$  (initial position). The angle  $\theta$  represents the variation from centre of field (initial position towards right or left). Current position is represented by points B and C.  $\theta$  is calculated using vector calculation.

$$\theta = \cos^{-1}[\langle \mathbf{V}_1, \mathbf{V}_2 \rangle / |\mathbf{V}_1| |\mathbf{V}_2|]$$

#### CONTROL FLOW

Point 1: When,  $\theta = (-20^\circ, 20^\circ)$  The robot moves forward as d is greater than 1.5m, and backward as less than 1.2m, and the robot stops as d ranges from 1.2m to 1.5m.

Point2: When,  $\theta = (20^\circ, 90^\circ)$  we should consider  $\theta$  firstly. The robot turns left until the value of  $\theta$  meeting the condition in Point 1 again, and then drive the robot as Point 1.

Point3: When,  $\theta = (-90, -20)$  we should turn the robot right until the value of  $\theta$  meeting the condition in Point 1 again, and then drive the robot as Point 1, too

## V. HARDWARE AND SOFTWARE REQUIREMENTS

- Hardware components
  - Kinect v2
  - Firebird v with ATMEL 2560 controller
  - USB 3 converter.
  - Laptop with windows 10.
- Software implementation
  - Software C#
  - Kinect SDK

## VI. APPLICATION

- Can assist in carrying loads for people working in hospitals, libraries, airports, etc.
- Can service people at shopping centers or public areas.
- Can assist elderly people, special children and babies.
- Can follow a particular person.

## VII. CONCLUSION

The robot is a three-wheeled robot with a Kinect sensor which can follow a human person without the need to wear a special device or clothing. In tracking the user, the robot uses the depth data from the Kinect as an input to the proportional-derivative algorithm to control the speed of the robot. By using depth estimation person at centre of field of the Kinect is detected and the user is followed. To avoid collision, the robot also uses the depth to check for obstacles that are between the user and the robot.

## VIII. RESULT

- With effort of Kinect detection accuracy improvement, the robot system is able to intelligently follow the user.
- Robot detects the position of person and calculate the distance using IR sensors and cameras.
- The object can be tracked even at the night as the Kinect has infrared camera which can identify the object at night.

## IX. REFERENCES

- [1] Praveen Manoharan, Nithin Sundararaj, Udhaya Sankar and S J Syed Ali Fathima in 2017: "REVIEW ON MICROSOFT KINECT'S GESTURE TECHNOLOGY BASED ROBOTICS".
- [2] S.A. Abdul Shukor\*, Mohammad Amir Abdul Rahim, B. Ilias: "Scene Parameters Analysis of Skeleton-Based Human Detection for a Mobile Robot using Kinect" Intelligent Human-Machine Systems and Cybernetics (IHMSC), 2016 4th International Conference
- [3] Guangzhu Peng, Chengguang Yang\*, Yiming Jiang, Long Cheng and Peidong Liang "Teleoperation Control of Baxter Robot based on Human Motion Capture". in International Journal of Control and Automation, vol. 7, no. 3, 2016, pp. 113–124.
- [4] Dajun Zhou, Fei Chao, Zuyuan Zhu, Chih-Min Lin, and Changle Zhou in 2016: "A Novel Approach to a Mobile Robot via Multiple Human Body Postures" in SICE Annual Conference (SICE), 2016 Proceedings of, 2015, pp. 2207 – 2211.
- [5] [5]Qimin Ren1, Qingjie Zhao1 , Hui Qi1, Lingrui "Real-time Target Tracking System for Person-following Robot".2016, IIT Kanpur, 2015
- [6] Robert Tubman and Khalid Mahmood Arif "Efficient People Search and Gesture Recognition using a Kinect Interfaced Mobile Robot", 2015 IEEE, pp. 1087-8270.
- [7] Yoonseon Oh and Songhwai Oh "Multiple-Hypothesis Chance-Constrained Target Tracking Under Identity Uncertainty" 2016 in Industrial Robot: An International Journal, vol. 32, issue 6, 2010.
- [8] Amruta Vinod Gulalkari, Giang Hoang, Pandu Sandi Pratama, Hak Kyong Kim, Sang Bong Kim , "Object Following Control of Six-legged Robot Using Kinect Camera".in International Journal of Advanced Robotic Systems, vol.208 ,2016, pp. 626-630.
- [9] Liying Cheng, Qi Sun , Han Su, Yang Cong , "Design and implementation of human-robot interactive demonstration system based on Kinect," in Control and Decision Conference (CCDC), 2012 24th Chinese, 2015, pp. 971-975.
- [10] Xuedong Huang, Acero. A, Adcock. J, Hsiao-Wuen Hon, Goldsmith. J, Jingsong Liu, Plumble. M, "Whistler: a trainable text-to-speech system," in Spoken Language, 1996. ICSLP 96. Proceedings, Fourth International Conference on, vol. 4, 2015, pp. 2387 - 2390