Experiment with Humanoid Robot Hand to Reach Object by Measuring Objects 3D Coordinates using Binocular Stereo Vision

Noushad Sojib
North East University Bangladesh
Sylhet-3100, Bangladesh

Sheikh Nabil Mohammad
Shahjalal University of Science & Technology
Sylhet-3114, Bangladesh

ABSTRACT

Stereo vision system adds the essential feature for robots to see the real world in a human-like manner by combining two 2D imaging systems. Robotic arm with end-effector helps robots to interact with real-world things by grabbing objects. Here in this research a 5 DOF human-like robot arm and a stereo vision set using two cameras mounted in parallel with 6cm distance was developed. Inverse kinematics is then calculated for the designed arm thus the robot can control the end-effector (gripper) position by adjusting motors angle. A software system was developed so the robot can perceive an objects 3d position using the stereo set and move the gripper through the help of kinematics.OpenCV blob detection technique was used to identify objects in an image. Summing up them the robot can now grip object seeing it in front of its stereo eye.

General Terms

Binocular Vision, Robot Control

Keywords

Humanoid Arm, Stereo Vision, Kinematics, Robot Vision, Epipolar Geometry

1. INTRODUCTION

The goal of this research is to develop a system to calculate near distance objects 3D coordinate with the help of stereo vision instead of active 3D scanner like LIDAR. The 3D coordinates then help the robot to pick up near distance object using its robotic hand. Robot sees the world using the camera and so its the field of computer vision that takes the responsibility to feed visual perception to the robot. The computer vision consists of image processing, machine learning, pattern recognition, etc. On the low label, it captures an image which is just a 2D representation of the 3D world & from that it could only perceive 2D information about objects. The single-camera vision system is unable to satisfy complete visual perception as it can not provide the depth of objects. Here comes the multiple camera stereo vision system to perceive the 3D nature of visual objects. NASAs solar observation system STEREO[1] is such an example of a stereo vision system.
In this research, two distinct topics are combined to work as a useful system for the robot. The first is finding objects 3D position and the second is moving the end effector accordingly to pick up target object. In the first part illustrates a brief description of finding objects 3d position, and the second part discuss the arm design, inverse kinematics[2], gripper sensor, etc. Lastly, the result of object detection and 3D coordinate calculation are compared with the Kinect sensor[3] to illustrates how well the system perform.

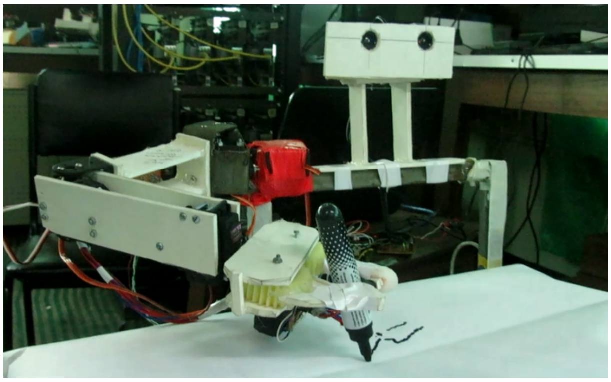

Fig. 1. Robot stereo vision and arm combination.

2. RELATED WORKS

Human eyes are capable of detecting difference in relative position when two objects sits in distinct distance and this is the disparity that helps brain to estimate objects distance[4]. Loop et al. showed how two 2D image of a 3D objects from two distinct viewpoints helps to calculate objects 3D position with the help of epipolar geometry.[5] This often subject to errors due to 2D capture of the 3D object[6]. Additionally, it require a certain of computation support to find the correlation of pixels between the pixels from the two image[7]. Visual servoing allows humanoid robots arm to manipulate objects[8] and such an experiment was done with Nao[9] robot to grab known object and planning motion[10].

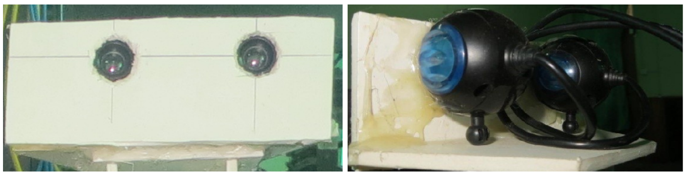

Fig. 2. Binocular Stereo Vision Setup.

3. HARDWARE SETUP

3.1 Binocular Vision

A stereo camera set was developed by placing two identical cameras mounted in parallel with a reasonable distance. To make it look human-like the cameras are kept in 6cm distance from each other.

3.2 5 DOF Humanoid Arm

A four DOF manipulator and a one DOF two-finger grippe were designed to work as the robot hand. The gripper contains two force sensor in fingers so it can understand when it gets in touch with object. The arm made of five servo motors which are connected with an Arduino-mega controller board and there are two 6v regulator to supply the necessary power to the motors.

3.2.1 Degrees of Freedom (DOF). DOF is the measurement of the manipulator which indicates how many directions the system can rotate. Generally, each motor creates one Degree of Freedom.

3.2.2 Link & Joints. Joints are the connection point that generally contains a single motor. Link is the physical media between two joints. The Table 1. shows the link dimension.

Table 1. Links Dimension

Link	Length(mm)
Shoulder	0
UpperArm (L0)	170
LowerArm (L1)	155
Gripper (L2)	105

3.2.3 1 DOF Gripper. This kind of two finger-gripper is simple and capable of grabbing almost every object and adding another finger adds huge extra complexity in control. There are two force sensor mounted on the fingertip.
The fingers were made of a hard PVC board with soft foam on the grabbing side. When it touches objects the controller sense increase force and then stop pressuring on the finger.

3.3 Force Sensor with gripper

Force sensors are used to perceive pressure. It changes resistance linearly with pressure and this change can be read using Arduino analog pin.

4. MEASURING OBJECTS 3D COORDINATE

4.1 Camera Calibration

Calibration is the first and the most important step of the stereo vision system. Performance mostly depends on it and it individually needs to be calculated for the individual stereo set. Here the calibration process was done using a chessboard which is a very

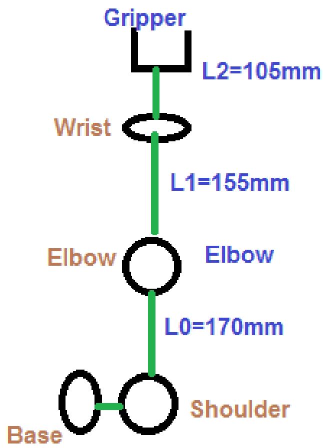

Fig. 3. 5 DOF robot Arm including end-effector

common technique as this makes it easier to find out sharp corners in an image.
A 9 X 6 sized chessboard of 2.7mm square was used to calibrate the cameras. From the calibration the calculated base line length =72mm , focal length= 763.7065mm and the projection matrix for our stereo set is given bellow.

1	0	0	-126.099
0	1	0	-279.856
0	0	0	763.705
0	0	0.422	-0.000

4.2 Rectification

Rectification is the process of aligning contents of two images. Thus we can calculate epipolar[11] geometry.

4.3 Stereo correspondence

Stereo correspondence is the most challenging part of the stereo vision system. There are several techniques to find correspondence in the right image what is it in the left image. Here we used semi block matching system for stereo matching. This method is fast enough to calculate in real-time. Stereo correspondence helps to calculate the disparity image from the left and right image.

4.4 Triangulation

After finding the disparity image the distance of the object is calculated with the help of camera intrinsic and extrinsic data calculated in the calibration process.
Simple steps to calculate disparity.

(1) We have two images.
(2) Find out P in left image, and where is it in right image.
(3) Disparity (d) = x position in left image (xl) x position in right image (xr)
(4) Z = b*f / d
(5) X = xl * Z / f
(6) Y = yl * Z /f

Fig. 4 shows the disparity map calculated.

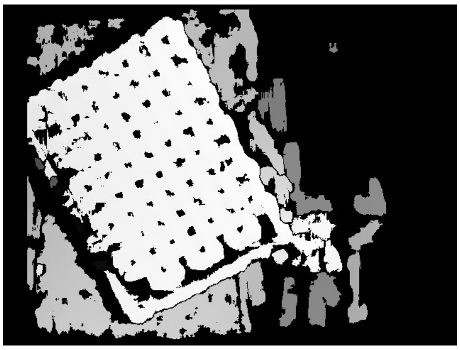

Fig. 4. Disparity Map

5. ROBOT ARM CONTROL

Arm control consists of two parts

(1) Lower level driver circuit & controller board Here two 6v regulator circuit were used to power all 5 motors. An Arduino-Mega is used as the low level motor controller and programmed in C++ language. This Arduino communicate with the attached Laptop using Serial Port & receive commands for individual motors as string data structure.

(2) High level kinematics control Kinematics is the most essential part for making the arm automatic. It requires a complex calculation of jacobian matrices or Cyclic Coordinate Descent (CCD) [12] to move end effector position precisely. Here a simple geometric 3d model of the arm was developed and its inverse kinematics was calculated therefore. The geometric model is less complex than the Jacobian and also works well found from the test experimented in real time.

5.1 Arm Kinematics

Generally, Jacobian[13] matrices and CCD algorithm are used for the calculation of arm kinematics. Calculation of these are most complex. Here authors made a simple algebraic solution followed by some trigonometric solution. The idea is simple, first rotate the base joint according to objects 3d position and then there left only a 2D plane. For the rest, first calculate object distance and apply it along with other links in a trigonometric problem. Fig. 5 shows the geometric representation of calculating kinematics.

6. EXPERIMENT AND RESULT

A tennis ball in red color was used to test the system. The two images from the binocular set were used to detect the object 2D position and construct the disparity map. From the disparity map, the distance from the robots center axis was triangulated to calculate the 3D coordinate of the ball. Fig. 6 shows the robot grabbing the red ball during an experiment. Authors have manually measured the calculated distance using measurement tape and found more error for the objects placed in larger distance. Additionally, authors calculated distance putting the Kinect sensor in the same position and results are compared for several measurements putting the object closer to far. The comparison is illustrated in the bar diagram Fig. 7 with 3 measurements. Distances in y-axis are in centimeter. The bar graph illustrates that the stereo set works near perfect for near distance (around 50cm) and the performance degraded much for higher distance. Besides this object 3D position finding authors also experimented how the robot hand approach towards objects or any target 3D position.

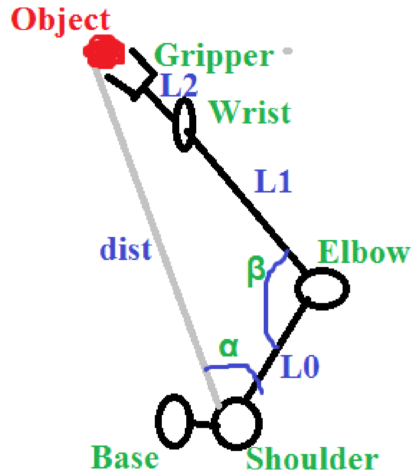

Fig. 5. Angle calculation for shoulder motor and wrist motor after rotating base motor.

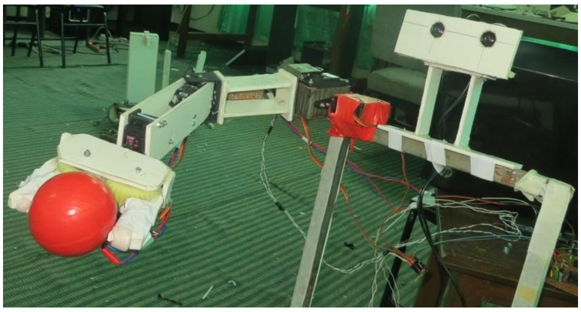

Fig. 6. The robot grabs a ball

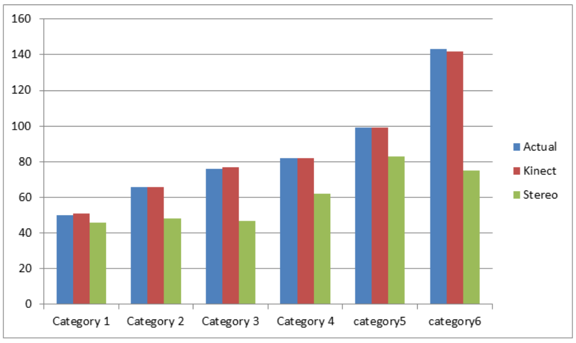

Fig. 7. Distance compared with Kinect sensor

7. CONCLUSION

During this work, the authors have made the robot arm, gripper and calculated kinematics for the robot in a simpler way. The robot system can now pick-up objects or draw simple shape by grabbing a marker in hand. Here authors have faced some difficulties in automatically object gripping as binocular based stereo vision system did not always show good results. From observation, the authors decided that in this problem active laser range finder could be a better choice to let the robot triangulate near distance object and pick the object automatically.

8. REFERENCES

[1] Socker, Dennis George, et al. ”NASA Solar Terrestrial Relations Observatory (STEREO) mission heliospheric imager.” Instrumentation for UV/EUV Astronomy and Solar Missions. Vol. 4139. International Society for Optics and Photonics, 2000.

[2] Asfour, Tamim, and Rdiger Dillmann. ”Human-like motion of a humanoid robot arm based on a closed-form solution of the inverse kinematics problem.” Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003)(Cat. No. 03CH37453). Vol. 2. IEEE, 2003.

[3] Smisek, Jan, Michal Jancosek, and Tomas Pajdla. ”3D with Kinect.” Consumer depth cameras for computer vision. Springer, London, 2013. 3-25.

[4] Marr, David, and Tomaso Poggio. ”A computational theory of human stereo vision.” Proceedings of the Royal Society of London. Series B. Biological Sciences 204.1156 (1979): 301-328.

[5] Loop, Charles, and Zhengyou Zhang. ”Computing rectifying homographies for stereo vision.” Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149). Vol. 1. IEEE, 1999.

[6] Murray, Don, and James J. Little. ”Using real-time stereo vision for mobile robot navigation.” autonomous robots 8.2 (2000): 161-171.

[7] Hirschmller, Heiko, Peter R. Innocent, and Jon Garibaldi. ”Real-time correlation-based stereo vision with reduced border errors.” International Journal of Computer Vision 47.1-3 (2002): 229-246.

[8] Claudio, Giovanni, Fabien Spindler, and Franois Chaumette. ”Vision-based manipulation with the humanoid robot Romeo.” 2016 IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids). IEEE, 2016.

[9] Gouaillier, David, et al. ”Mechatronic design of NAO humanoid.” 2009 IEEE International Conference on Robotics and Automation. IEEE, 2009.

[10] Mller, Judith, Udo Frese, and Thomas Rfer. ”Grab a mug-object detection and grasp motion planning with the Nao robot.” 2012 12th IEEE-RAS International Conference on Humanoid Robots (Humanoids 2012). IEEE, 2012.

[11] Zhang, Zhengyou, et al. ”A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry.” Artificial intelligence 78.1-2 (1995): 87-119.

[12] Saha, Ankan, and Ambuj Tewari. ”On the nonasymptotic convergence of cyclic coordinate descent methods.” SIAM Journal on Optimization 23.1 (2013): 576-601.

[13] Buss, Samuel R. ”Introduction to inverse kinematics with jacobian transpose, pseudoinverse and damped least squares methods.” IEEE Journal of Robotics and Automation 17.1-19 (2004): 16.