Automated Monocular Vision based System for picking Textureless Objects

Tanya Raghuvanshi
Electrical Engineering
Indian Institute of Technology
New Delhi

Shraddha Chaudhary
Electrical Engineering
Indian Institute of Technology
New Delhi

Varnica Jain
Electronics Engineering
Temasek Polytechnic
Singapore

Sumeet Agarwal
Electrical Engineering
Indian Institute of Technology
New Delhi

Santanu Chaudhury
Electrical Engineering
Indian Institute of Technology
New Delhi

ABSTRACT

This paper proposes an Autonomous Machine Vision system which grasps a textureless object from a clutter in a single plane, rearranges it for proper placement and then places it using vision. It contributes to a unique vision-based pose estimation algorithm, collision free path planning and dynamic Change-Over algorithm for final placement.

1. INTRODUCTION

In many manufacturing environments, work pieces lying in a bin are hand-picked and then assembled on assembly lines. The work environment in such manufacturing areas is generally unhealthy and harmful. Human limbs are exposed to danger as work-pieces are often inserted into the machines by hands. Such jobs are monotonous and do not enrich human life. Hence, there is a dire need for the automation of process, which proves most fruitful and advantageous considering the safety, accuracy and time reduction through Computer Vision.

Deriving inspiration from human vision, the pose of a workspace can also be estimated and processed through computer vision. The information obtained, can further be transformed into action commands for controlling machines and for performing the required assembly. Some of the known approaches for such a pose estimation include use of Stereo Vision[1] and range sensors[2]. In a stereo set-up, the correspondence between different views can be used to reconstruct a scene. However, such a reconstruction requires atleast one point of correspondence, which is generally, a unique permanent point in the view or features of the object. Permanent correspondence points requires a great precision in calibration and makes work-space sensitive to small disturbances causing instability for use in a dynamic industrial environment. Therefore, using features as correspondence points becomes vital for study of pose and the object. Most popular feature extraction techniques are based on gradient of the object(SIFT [3][4], SURF, colour thresholding). However, in practical industrial use,objects are found to be idealistic and the gradient-based methods are bound for failure on such texture-less work-objects. Moreover, feature extraction algorithms can, in turn, add to the extra cost and time.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

AIR '15, July 02 - 04, 2015, Goa, India

© 2015 ACM. ISBN 978-1-4503-3356-6/15/07…$15.00

DOI: http://dx.doi.org/10.1145/2783449.2783514

1.1 Related Work

Bin-picking invokes several problems at every stage: data acquisition, pose estimation, grasping, arrangement, collision avoidance[9] could be few a of them. Research on Collision free grasping[8], describes probability estimation technique for grasping an object.Another research [11] uses structured light to determine the pose of an object in workspace.

There has been a lot of significant work striving to achieve collision free grasping. However, an independent, robust and time efficient system is still unavailable.Also, the existing solutions, make use of stereo set-up. In stereo set-up, two or more cameras are used to generate 3D reconstruction of a work-space and different poses of objects are estimated. For 3D reconstruction, correspondence between two images of a given scene is found to be matching at some fixed points on the objects, and is common in both images. Moreover, existing system solutions for bin-picking do not deal with texture less objects or use mono vision solution. Finding correspondence points in a texture less or identical object is very difficult and even impossible in certain cases. On the other hand, humans can easily observe and detect pose of objects in a clutter, even in 2D pictures. Hence, in this paper, we propose an approach that uses mono-vision model based recognition to solve pose estimation problem for occluded objects. The previous research [10], in the area, follows a database-based approach. Pellets (texture less, identical work-objects for study), were simulated before-hand and their possible orientations and positions were stored in a database. These simulations were run for discrete positions and discrete orientations of the pellets. A contour-based method was then implemented, to identify the actual pellet orientation. This required a significantly long duration of online execution time and also a memory to store the database. Therefore, an essential need was felt, to develop a better algorithm or technique.

In comparison with the state of the art algorithms, the

proposed technique would require less memory and processor power, which in turn, would make the entire system cost effective and easy to implement. We propose a mono-vision model based recognition, to do pose estimation of a feature less object in a jumbled configuration, for the final placement or arrangement. Paper is organized as follows - Section 2, discusses the brief overview of the system; Section 3 discusses novelty in the position estimation of the isolated and occluded pellets, along with its novel path planning algorithm for the placement of the pellets; Section 4 discusses results; Section 5 presents the contribution, comparison and conclusions.

2. METHODOLOGY :BRIEF OVERVIEW OF THE SYSTEM

Our goal is to pick-up pellets from a bin and then finally place or arrange them. For the purpose of this research, KUKA robotic arm was used for all the experiments. It is equipped with a Gripper, with two fingers and a camera, which is mounted at the end of the arm. This camera is used to capture all images. The objects are black cylinders/pellets with no sharp edges or any peculiar texture. The pellet is extracted from the image acquired by the camera to be picked up, and placed at a Change-Over position, for rearrangement to be picked-up again, for final insertion into a horizontal pipe. Experiments with pellet placement in grid and bin dropping were also performed and are discussed in this paper.

Algorithm 1 Brief overview

Step 1: CALIBRATION:calibrate the workspace using Zhang's calibration [6].
Step 2: SEGMENTATION: Segment the Object of interest using S.K. Nayar's Segmentation technique [5].
Step 3.1: Position estimation of the Isolated Pellets.

- Using Shape based Approach.
- Using machine learning Approach.

Step 3.2: Position estimation of Occluded Pellets in clutter using Heuristic Approach.
Step 4: Path Planning and Rearrangement.
Step 5: Hole Insertion/Pellet Placement/Bin dropping.

Pellets are observed by camera in one of the following configurations: Vertical/ Standing; Horizontal/ Sleeping or Inclined/ Slant. Standing pellet rests on circular base of cylinder(fig 1), Sleeping pellet rests along the height of the cylinder(fig 2) and Slant pellet rests with one edge on another pellet(fig 3).Conventional, image processing techniques which are based on color thresholding, fail to extract a pellet from a cluster of pellets, since such techniques cannot distinguish an object from its own shadow, which is commonly observed in our set-up.. Other gradient-based techniques are also fruitless, as the pellets do not have any sharp edges and are hence, featureless with no texture. Therefore, segmentation proposed by Nayar [5] was implemented for the purpose.

S.K. Nayar [5] proposed segmentation on the basis of a material's reflectance ration property. It is the measure of light an object can reflect. Every object has different reflectance ration property. Since the reflectance ration for pellets is not known, relative reflectance ration is used to

Figure 1: Standing pellet

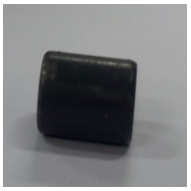

Figure 2: Sleeping pellet

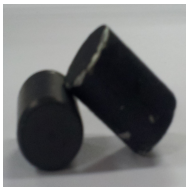

Figure 3: Slant pellet

isolate different pellets. Briefly, the method finds relative reflectance ratio for each pixel by comparing the intensity value of its neighboring pixels. Fig 13 and Fig 14 show segmentation using Nayar's Approach under high illumination conditions.

After segmenting the image, contours around each segmented pellet are found. These contours describe the periphery of each pellet. A bounding square is calculated for each of these contours. This cubical area is then used to extract the pose of the pellet. Pose consists of pellets position (X,Y,Z - in world) and its orientation i.e. inclination w.r.t to its vertical plane. To find the world coordinates of the pellet, center of each extracted pellet is calculated. Thereafter, the calculated image coordinates of the center are then converted to world coordinates using Zhang's calibration[6].

Zhengyou Zhang proposed camera calibration[6] which is one of the known techniques used. In this methodology, a planar pattern is observed from different orientations (at least two). Equation 1 gives Zhang's calibration equation.

$$c*\begin{bmatrix}x\\y\\1\end{bmatrix} = \begin{bmatrix}R_{11} & R_{12} & t_1\\ R_{21} & R_{22} & t_2\\ R_{31} & R_{32} & t_3\end{bmatrix}^{-1} * \left(\left(K_{(3X3)}\right)^{-1} * \begin{bmatrix}u\\v\\1\end{bmatrix}\right) - Z* \begin{bmatrix}R_{13}\\R_{23}\\R_{33}\end{bmatrix}$$

Equation 1: Zhang's Calibration Equation

3. POSE ESTIMATION AND PLACEMENT

In this section, an end to end system is described, which unfolds the steps mentioned in Algorithm 1 and proposes a complete solution to the described Pick-up problem.

3.1 Position estimation of Isolated pellets

After segmenting out the pellets and finding their position using calibration, orientation of a pellet is estimated from the extracted cubical area. Standing pellets are detected using Hough Transform as shown in fig: 5. Considering the symmetry of the object, we used Singular Value Decomposition(SVD) for orientation calculation of Sleeping pellets.

3.1.1 Shape based approach

Singular Value Decomposition(SVD) factorizes the input covariance matrix into left and right singular values and a diagonal singular matrix. The right singular matrix consists of the Eigen vector which gives the orientation of pellet(Fig: 4).

Although SVD of pellets provided good results, it is still dependent on the quality of segmentation performed on the raw image. It fails, when shadow of the pellet is accidentally segmented with the pellets. The accuracy of orientation calculated decreases on changing the view point of the robot. It is dependent on the orthogonality of the camera with respect to the plane of pellet pick-up. Therefore, faster and

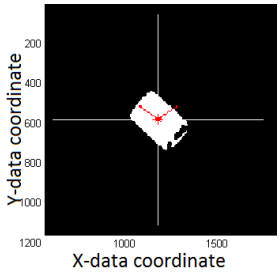

Figure 4: SVD orientation

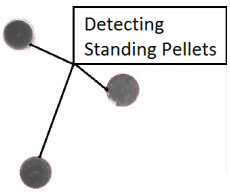

Figure 5: Standing Pellet Detection using Hough Circles

more accurate technique like Machine Learning was used.

3.1.2 *Using Machine Learning Approach*

Machine learning is a branch which studies algorithms and it's construction by learning from data. It constructs models from data provided. This data is in the form of features which describe the object optimally. Pellets are defined by their height and radius which indicates the presence of feature descriptors which are independent of texture. Although features can be extracted for non-symmetrical objects by using well known algorithms like gradient based methods, but for symmetrical objects like pellets, features can only be extracted using non-gradient based techniques like Principal Component Analysis(PCA). PCA is based on finding the direction which is a linear combination of all features and captures the maximum variance in the database. The first component captures the highest variance, second component the second highest and so on. The Principal Components for dataset of 1.5K images are shown in fig: 6(a,b,c,d). As seen, maximum variance is observed for first and second components. Third and Forth principal components are harmonics of the previous two components.

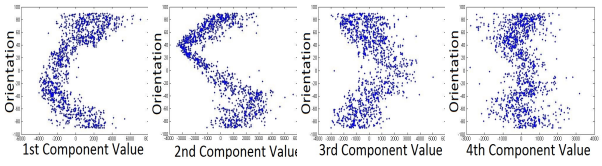

Figure 6: (a,b,c,d) 1st, 2nd, 3rd, 4th Principal Components

The behavior is captured by training a Support Vector Regressor(SVR) model for principal component values for each orientation of the object. Results for Non-linear kernel are shown in fig: 7 and fig: 8.

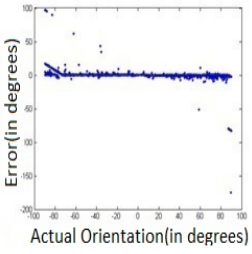

Figure 7: Actual-Error Analysis

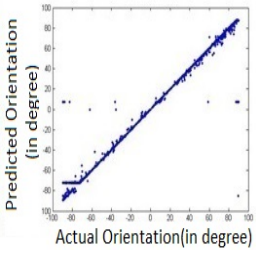

Figure 8: Actual-Prediction Analysis

3.2 Position Estimation of Occluded pellets

In this section, we present the position estimation of Occluded pellets as shown in Fig: 12, the presented approaches (SVD and Machine learning algorithms) for isolated pellets were based on the assumption that area extracted had only a single pellet. Area extraction for pellet detection is based on segmentation and hence partially occluded pellets cannot be retrieved properly for they are seen as a single area. Machine learning approach accounts for the change in all the pixel values of the extracted pellet. Since some of the feature/pixel weights are added up or are missing, the change is not captured and hence the prediction for orientation is incorrect. Therefore, using the segmented image, contours were then used for handling the occluded pellets scenario.

Segmentation separates out background and foreground effectively. Contours are then, extracted from each cluster. Each extracted contour is approximated to a polygon[7] and further processing is based on these approximated contours.

3.2.1 *Using Heuristic Approach*

Hypothesis for heuristic approach presented follows:
There will be at least one pellet in the segmented cluster which will always share it's longest edge along the outline of the contour extracted after image segmentation.

Hypothesis Validation: For implementation of this hypothesis, Longest edge from the approximated polygon is observed. Orientation of the pellet is assumed to be same as that of the Longest edge. Calculations were done by using one of the vertices of the Longest edge of the approximated polygon. Center of the pellet was calculated in two stages. In the first stage, a point on the edge near the vertex was calculated. In the second stage, depending on the pixel values around this point, center is calculated either by moving towards left or right of this point in the direction of orientation in order to position Center on the pellet. This hypothesis worked correctly for the scenario with only Sleeping pellets.

Algorithm 2 Algorithm for Heuristic Approach

1. Find contours, approx. to polygon & find longest edge.
2. Orientation of pellet same as that of the longest edge.
3. If the vertices of the longest edge observed was in approximate radius of the circle obtained(standing pellet), position is calculated using the center of the circle.
4. If the vertices do not belong to any standing pellet, then using one of these observed vertices, center of the pellet is calculated in two stages. In first stage, point on the edge near vertex is calculated assuming it be the longest edge of the pellet. In the second stage, depending on the pixel values around this point, center is calculated either by moving towards left or right of this point in the direction of orientation.

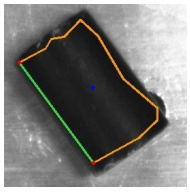

Figure 9: Results for sleeping pellets

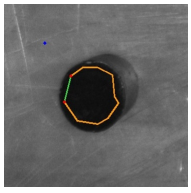

Figure 10: Wrong results

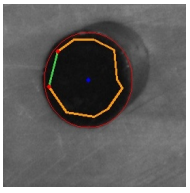

Figure 11: Results after hypothesis fix

In order to include more general cases where Standing pellets were also observed, the hypothesis was modified. The Longest edge of the contour could belong to a sleeping pellet or standing pellet. Standing pellets were detected using Hough Transform. Hence, by thresholding, we were able to identify if pellet segmented was Sleeping or Standing. If the vertices of the observed Longest edge were in approximate radius of circle obtained (standing pellet), position was calculated using the center of the circle. These corrections are shown in fig: 11 and 10. Orange contour represents the approximated polygon , green line represents the longest edge detected, red points on this edge are the vertices and blue point is the center calculated for the pellet.

View considered was orthographic, hence orientation and position of Inclined pellets would be same as to when calculated by the hypothesis stated above. Fig: 12a shows cases handled with Inclined pellets. The pellet detected is a Sleeping pellet. Fig: 12b shows Inclined pellet detected. Fig: 12c shows a Sleeping pellet detected with two Standing pellets.

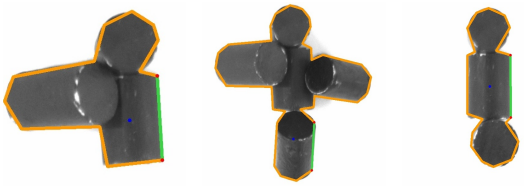

Figure 12: (a,b,c) Result Using Heuristic Approach

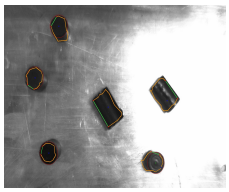

Figure 13: Final image with positions

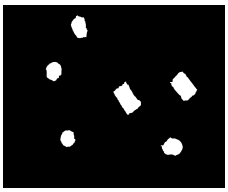

Figure 14: Segmentation of the same image

3.3 Path Planning and Rearrangement

This section deals with path planning for the pellet to be picked-up from occluded configuration. As shown in fig: 12 (a,b,c), pellets to be picked up are either surrounded by other pellets or are supporting a Slant pellet. The pellets which are restricting the pick-up are the Hurdle pellets, while the pellet to be picked up is the Selected pellet. If manipulator goes to pick-up the Selected pellet in occluded scenario, the gripper would hit the Hurdle pellets. This would cause damage not only to the pellet but also the manipulator. Hence, we proposed a path planning algorithm which clears the Hurdle pellets using the gripper and then picks up the Selected pellet. Hindrance by other pellets decides the path for the pellet to be picked-up by gripper. For the detected/chosen pellet, six sections are drawn around in the same orientation as that of the Selected pellet. The sections are numbered as shown in fig: 15. Each section is drawn as a yellow colored rectangle. The dimensions of these sections are same as the dimensions of the rectangular area covered by the gripper on the pick-up plane. One of the sides is shorter and referred as Shorter edge while the other side as Longer edge. The basic attempt of the algorithm is to push Hurdle pellets and then to approach the Selected pellet.

Sections which are considered to be neighbors of selected pellets are section 1, 2, 3 and 4.

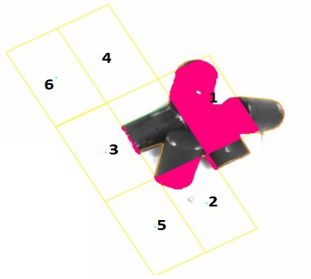

Figure 15: Nomenclature for sections around pellet

Algorithm 3 Path Planning

Case 1: Surrounded from none. **Path:** Center of pellet.
Case 2: Surrounded from one side.

- **Case 2a:** Either of the shorter edges.
- **Path:** Other empty region around the shorter edge of pellet, then to the center of the detected pellet. (if 1, then 3-> center & if 3, then 1-> center)
- ***Case 2b:** Either of the longer edges.
- **Path:** Nearer empty region around the shorter edge of pellet,then to the center of the detected pellet. (if 24, then nearer of (1 or 3)-> center).

Case 3: Surrounded from two sides

- ***Case 3a** One shorter edge other long edge.
- **Path:** Other empty region around shorter edge of pellet, then to center of pellet. (if 1 & 2or4, then 3-> center & if 3 & 2or4, then 1-> center)
- **Case 3b:** Both shorter edges of the pellet(1&3).
- **Path:** Center of the detected pellet. (if (2&4), then nearer->center).
- ***Case 3c:** Both longer edges of the pellet(2&4).
- **Path:** Nearer empty region of shorter edge of pellet, then to the center of the detected pellet. (if (2&4) then nearer-> center).

Case 4: Surrounded from all but one edge.

- ***Case 4a** Not from one of the shorter edges.
- **Path:** Empty region around the shorter edge, then to the center of the detected pellet.(if !1 then 1->center, if !3 then 3->center).
- ***Case 4b** Not from one of the longer edges.
- **Path:** Empty region around the longer edge of pellet and then take U-shape path from this region to region around nearer-shorter edge and then finally to the center of the detected pellet. (if !2 then 4->6->3->center, if !4 then 2->5->3->center).

Case 5: When surrounded from all edges of the pellet.
Path: Do not pickup this pellet, it will hit other pellets.
*** Removes hurdle pellets from path.**

3.3.1 *Rearrangement/Changeover*

Although, the above algorithm ensures isolation of a single pellet from a cluster of pellets, but pellet needs to be gripped upright for any kind of placement. Standing pellets are gripped upright and hence require no rearrangement. On the other hand, Sleeping and Slant pellets are taken to Change-over position for upright gripping.

The gripper's fingers are aligned horizontally with respect to the plane of the Change-over position. If the picked up pellet is an Inclined pellet, depending upon the magnitude of it's inclination, it will either be placed horizontally like a Sleeping pellet(16) or placed vertically like Standing pellet(fig: 18). On the other hand, Sleeping pellet will always be placed vertically. Vertical and horizontally kept pellets are differentiated using Hough Circle detection. The horizontally placed pellet is picked-up in the same manner as it was placed. The gripper, positions parallel to the Change-over plane for the pick-up.(fig: 17) The vertically placed pellets are gripped upright. In this case, gripper positions itself perpendicular to the Change-over plane.(fig: 19)

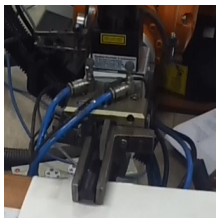

Figure 16: Placing for Standing Pellet

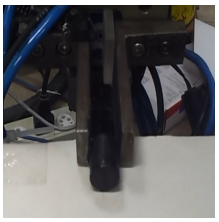

Figure 17: Pick-up for sleeping pellet

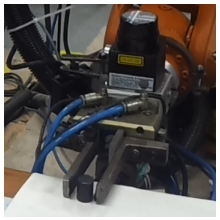

Figure 18: Placing for Sleeping Pellet

Figure 19: Pick-up for standing pellet

3.4 Placement

In this section, we discuss the final placement of the pellet. After it's successful rearrangement at Change-over position, experiments with three frequently used placements were performed: Bin-dropping, Grid making and Hole insertion.

3.4.1 **Bin dropping:**

Robotic manipulator interacts with another robot i.e.. two or more robots interact with each other simultaneously. A bin is kept within the reach of robotic manipulators. Every time robot picks up pellet, it goes to the bin and then drops it. This dropping point can be generalized to an interaction point of two robots. Figure 22 show bin dropping.

3.4.2 **Grid making:**

Robotic manipulator assembles the picked-up pellets in a tray which is portrayed as a representation of any assembly line and may require more than one object's placement. An example: to put toys in different boxes, while being kept on same the assembly line. A rectangular grid has been defined for this purpose. Everytime robot picks up a pellet, it goes to a new grid position and places it upright at the defined position. Figure 20 and 21 show grid making in process.

Figure 20: Grid Making with 2nd pellet

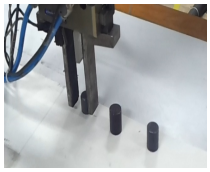

Figure 21: Grid making with 3rd pellet

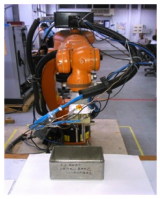

Figure 22: Bin dropping

3.4.3 **Hole insertion:**

It is needed when robotic manipulator requires to insert an object with precision into a small opening or a hole. For example: insert screws for a machine. Holes for screw insertion need to be precisely calculated. A similar condition was recreated for pellet Hole-insertion. Two pipes were mounted on an aluminum plate and every time pellet was to be inserted, an image was captured, where circles (pipe holes) were detected. Pellets were inserted, one-by-one, in the extracted hole positions. Fig: 23 shows holes detected for insertion. Fig: 24 and fig: 25 shows stages for hole insertion.

Figure 23: Holes detected using hough circles

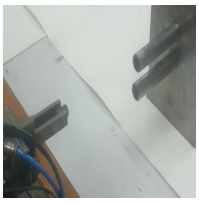

Figure 24: Grippers approach for pellet hole insertion

Figure 25: Pellet Insertion in Hole by Manipulator

4. RESULTS

This section presents results based on a number of trails run(on robotic arm) for various steps mention in section 3.

4.1 Multiple stages in pellet pick up

Every time a pellet is removed, presented algorithm calculates the pose of the next pellet to be picked up. Fig 26 shows different views observed as pellets are removed one by one by the robotic manipulator. It indicates the robustness and reliability of the manipulator and algorithm.

S.No	Algorithm Used	Accuracy(in %)
1	Pick-up efficiency	97.33%
2	Hole Insertion efficiency	94.6%

Table 1: Algorithm Accuracy

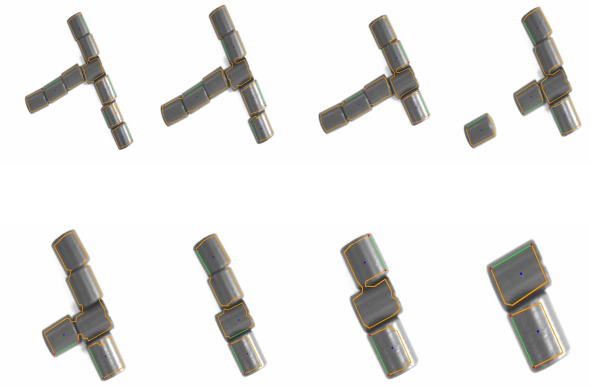

Figure 26: Multiple Stages in pellet pickup

4.2 Reliability of the Algorithm

As shown in fig: 27, because of variation in size of pellets accuracy for hole insertion is less. A smaller sized pellet may be gripped properly at Change-over position, however there will not be much clearance for pellet to be inserted in the hole completely. Hence, at the time of insertion although the manipulator is accurate, pellet insertion fails.

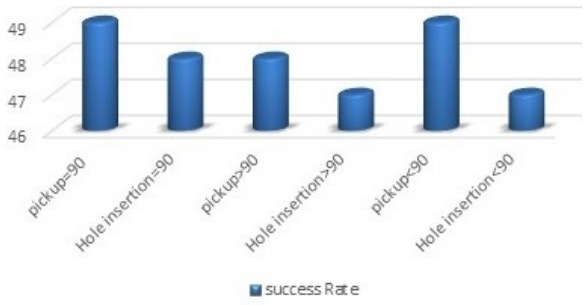

Figure 27: Success rate for hole insertion

4.3 Error in the estimated orientation

In this section, the average error to estimate orientation is calculated, which is 1.67 Degrees. Fig: 28 shows error observed to calculate Orientation using Heuristic approach.

-70	-30	-10	0	10	30	50	60	75	90
1.55	-1.12	-1.77	-2.51	2.23	-1.55	-1.01	2.14	1.11	-1.7

Figure 28: Error Plot for Orientation

5. DISCUSSION AND CONCLUSIONS

We have presented a complete Autonomous Vision System for both Isolated and Occluded pellets. For isolated pellets, Machine Learning approach is clearly better than previous approach which was based on Database Matching[10], hence the CPU Time and memory requirement has been reduced drastically. For occluded pellets, hypothesis can be validated for the expected configurations. It is also capable of handling isolated pellets with less time and computational complexity than the existing solutions to handle occluded object extraction. Heuristic based approach can handle not just Sleeping and Standing pellets but also Slant pellet. Single layer 3-D is easily handled with this research's Heuristic approach. Path planning provides a natural subconscious human approach to pick-up pellets. It successfully removes the unwanted pellets and poses an intelligent dynamic solution for selective picking. Pellet rearrangement asserts a novel mechanism which involves intelligent pose detection and final placement.

6. REFERENCES

[1] Hartley, R. and Zisserman, A. Multiple View Geometry in Computer Vision, Cambridge University Press, Cambridge 2001.

[2] Simon, Hebert, M. and Kanade T. 1993. Real-time 3-D Pose Estimation Using a High-Speed Range Sensor, Robotics Institute, Paper 460.

[3] Bjï£¡rn, J. and Moe, A. 2005. Patch-duplets for object recognition and pose estimation,Proceedings. 2nd Canadian Conference on Computer and Robot Vision, 9-16.

[4] Changchang, Wu. 2008. 3D model search and pose estimation from single images using VIP features, Computer Vision and Pattern Recognition Workshops, CVPRW'08, 1-7.

[5] S.K.Nayar, Bolle,"Computing reflectance ratios from an image",Pattern Recognition, 26(10):1529-1542, 1993

[6] Z. Zhang. "A flexible new technique for camera calibration", IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11):1330-1334, 2000

[7] David Douglas & Thomas Peucker, "Algorithms for the reduction of the number of points required to represent a digitized line or its caricature", The Canadian Cartographer 10(2), 112-122 (1973)

[8] Spenrath, Felix; Spiller, Alexander; Verl, Alexander, "Gripping Point Determination and Collision Prevention in a Bin- Picking application," Robotics; Proceedings of ROBOTIK 2012; 7th German Conference on , vol., no., pp.1,6, 21-22 May 2012

[9] Pochyly, A.; Kubela, T.; Singule, V.; Cihak, P., "3D vision systems for industrial bin-picking applications," MECHATRONIKA, 2012 15th International Symposium , vol., no., pp.1,6, 5-7 Dec. 2012

[10] P. Tiwan, Riby A. Boby, Sumantra D. Roy, S. Chaudhury, S.K.Saha,"Cylindrical Pellet Pose Estimation in Clutter using a Single Robot Mounted Camera",2013 ACM July 04-06 2013

[11] Jong-Kyu Oh; KyeongKeun Baek; Daesik Kim; Sukhan Lee, "Development of structured light based bin picking system using primitive models," Assembly and Manufacturing, 2009. ISAM 2009. IEEE International Symposium on , vol., no., pp.46,52, 17-20 Nov. 2009