

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



# Non Contact 2D and 3D Shape Recognition by Vision System for Robotic Prehension

Bikash Bepari, Ranjit Ray and Subhasis Bhaumik  
*Haldia Institute of Technology, Haldia, India*

*Central Mechanical Engineering Research Institute, Durgapur, India*  
*Bengal Engineering and Science University, Shibpur, Howrah, India*

## 1. Introduction

The increasing demand for robotic applications in today's manufacturing environment motivates the scientists towards the design of dexterous end-effectors, which can cope with a wide variety of tasks. The human hand can serve as a model for a robotic hand for prehension, which interacts with the environment with more than twenty-seven degrees of freedom (dof<sup>s</sup>). The articulation and accommodation of the human hand in terms of dexterity, stability, tactile and /or nontactile sensation, stiffness, weight to grip force ratio, resistance to slip, and adaptability provide clear direction for active research in dexterous manipulation by the robot.

Robotic hands, which have been developed over past three decades, sometimes on adhoc basis to accomplish a predefined task. In robotic prehension, perhaps the choice of grip to ensure stable grasp has become the main hurdles to confront with. Since the inception of development of robotic hand, this has been revealed from thorough research that, more emphasis has been given on the electro-mechanical considerations than on the stability of the grasp, that too under vision assistance. This leads to an effort to develop new algorithms for stable grasp irrespective of its shape.

Prehension combines the choice of grip and the act of grasping including its control. A desire, internally or externally generated, triggers responses in the memory and the eyes; the hand instantaneously takes a position over the object, grips it and manipulates the task. The process of prehension is controlled by feedback loops with the position and the force of the fingers as shown in Fig.1. The eye is basically a sensor, whereas the brain is the central processing unit, which sends signal to the different muscles and tendons to act accordingly to actuate the hands for manipulation. Choosing a grip is an important step in the process of prehension. The block diagram in Fig. 2 illustrates the relationship between the different activities included in prehension. This can be simulated by using a vision system on a robot and a processor with feedback loop to manipulate the tasks by an appropriate gripper in the similar pattern of eye-memory analysis as shown in Fig. 3.

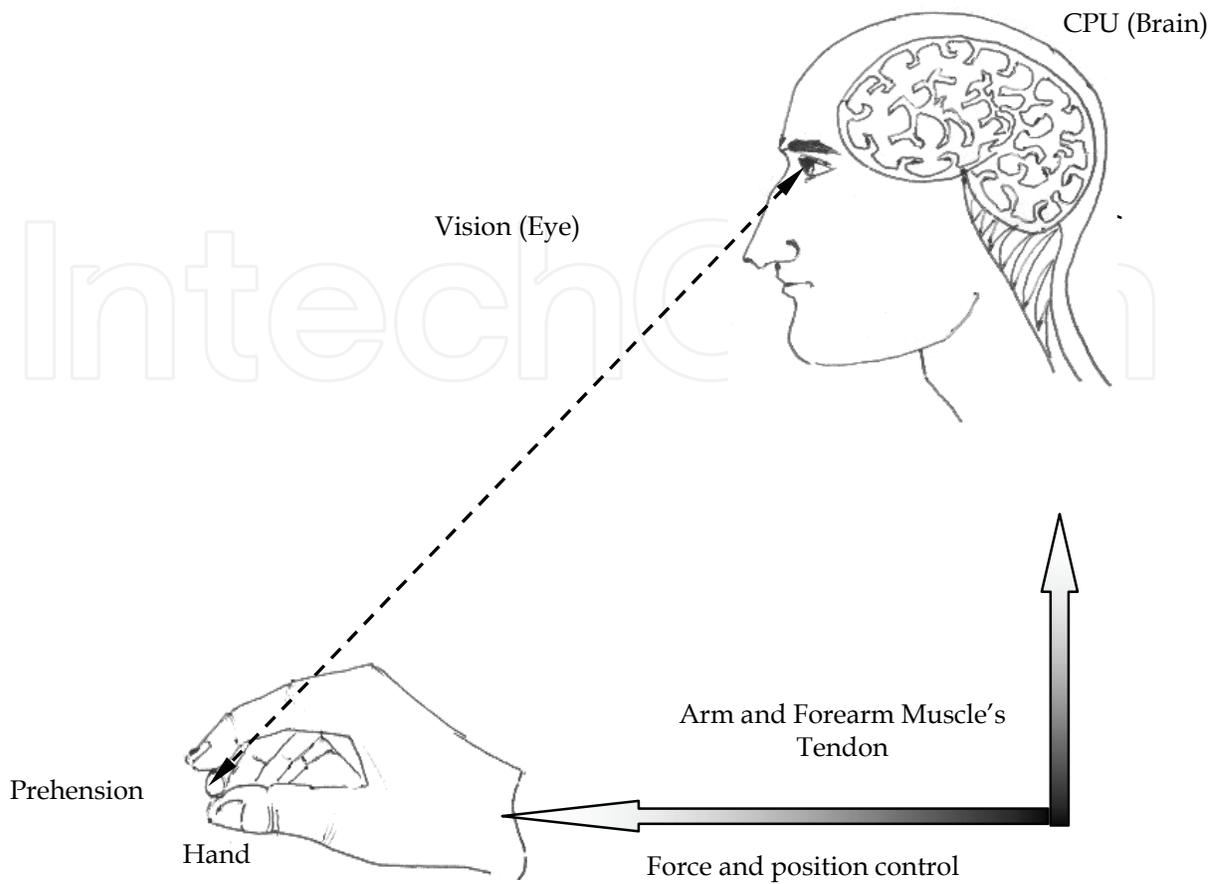


Fig. 1. The eye -memory integration-based human prehension

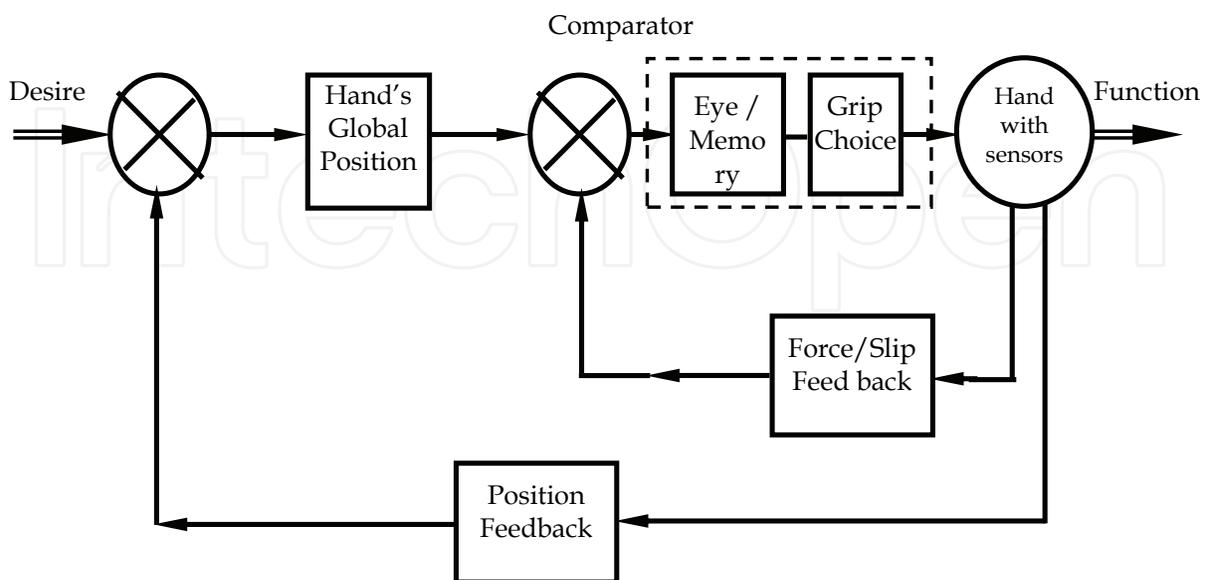


Fig. 2. Feedback diagram of the actions involved in prehension

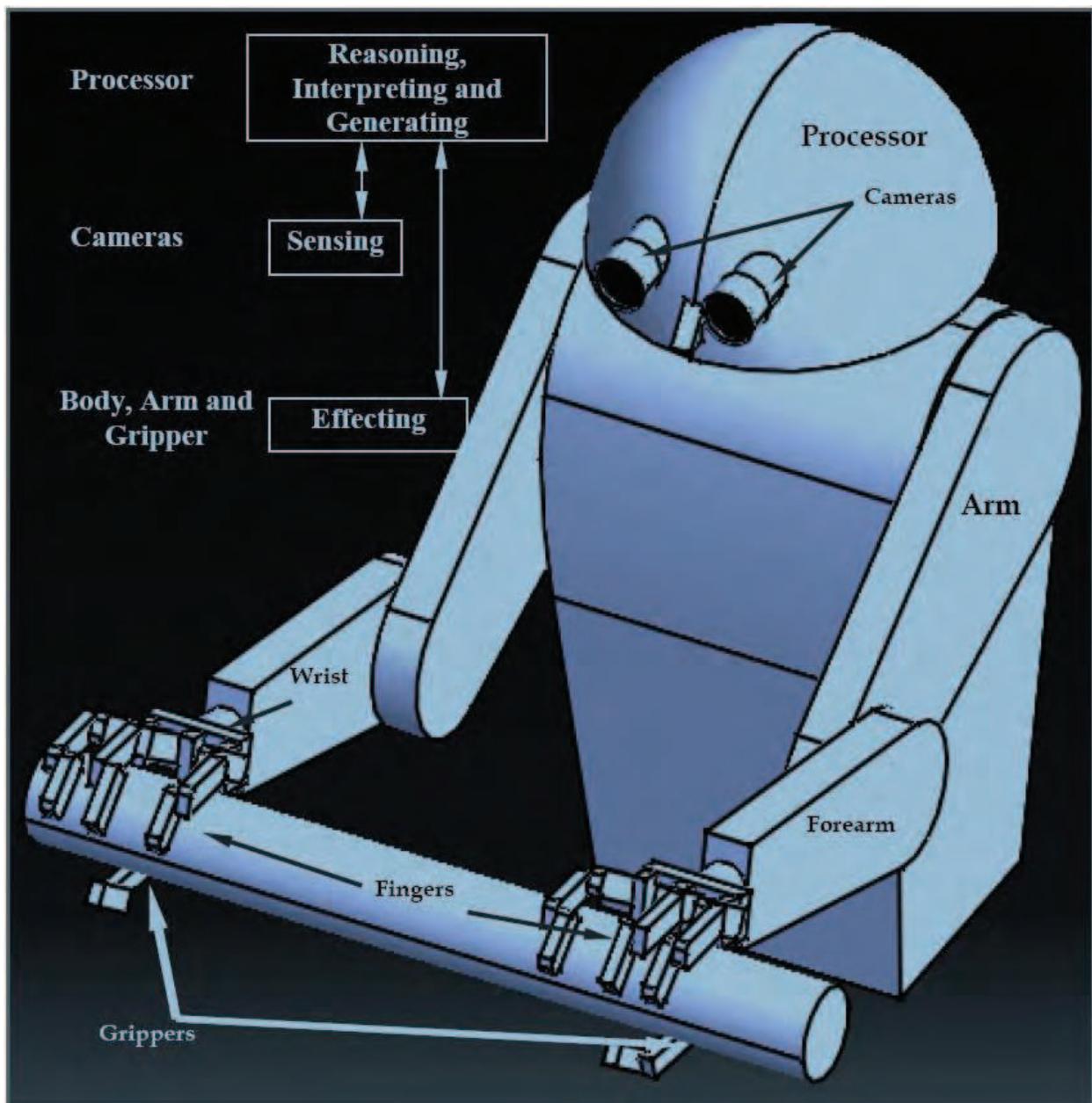


Fig. 3. Simulation of the robotic prehension

The concept of replication of the attributes of human morphology, sensory system and neurological apparatus alongwith the behavior leads to a notion of embodiment - this in turn over time is refined, as the brain and physiology change. If the grasping modes are critically examined between a child and an adult, a distinguishable difference may be observed between the two. The presense of simultaneous path planning and preshaping differentiates the latter one from the former.

The essential goal of the present work was to ensure the stability of a grasp under visual guidance from a set of self generating alternatives by means of logical ability to impart adroitness to the system (robotic hand).

## 2. Survey of Previous Research

As a step towards prehension through vision assistance in robotic system, Geisler (Geisler, 1982) described a vision system consisting of a TV camera, digital image storage and a mini computer for shape and position recognition of industrial parts.

Marr (Marr, 1982) described 3-D vision as a 3-D object reconstruction task. The description of the 3D shape is to be generated in a co-ordinate system independent of the viewer. He ensures that the complexity of the 3-D vision task dictates a sequence of steps refining descriptions of the geometry of the visible surfaces. The requirements are to find out the pixels of the image, then to move from pixels to surface delineation, then to surface orientation and finally to a full 3-D description.

Faugeras (Faugeras, 1993) established a simple technique for single camera calibration from a known scene. A set of ' $n$ ' non-co-planar points lies in the 3-D world and the corresponding 2-D image points are known. The correspondence between a 3-D scene and 2-D image point provides an equation. The solution so obtained, solves an over-determined system of linear equations. But the main disadvantage of this approach is that the scene must be known, for which special calibration objects are often used. Camera calibration can also be done from an unknown scene. At least two views are needed, and it is assumed that the intrinsic parameters of the camera do not change.

Different researchers like (Horaud et al., 1995), (Hartley, 1994) and (Pajdha & Hlavac, 1999) worked on this approach. Horaud considered both rotational and translational motion of the camera from one view to another. Hartley restricted the camera motion to pure rotation and Pajdha et al. used pure translational motion of camera to get linear solution. Sonka et al. (Sonka, 1998) discussed on the basic principle of stereo vision (with lateral camera model) consisting of three steps: a) Camera calibration, b) Establishing point correspondence between pairs of points from the left and the right image and c) Reconstruction of 3D coordinates of the points in the scene.

David Nitzan (Nitzan, 1988) used a suitable technique to obtain the range details for use in robot vision. The methodology followed in his work is, image formation, matching, camera calibration and determination of range or depth. Identifying the corresponding points in two images that are projections of the same entity is the key problem in 3D vision. There are different matching techniques for finding the corresponding points.

Victor et al. (Victor & Gunasekaran, 1993) used correlation formula in their work on stereovision technique. They determined the three-dimensional position of an object. Using the correlation formula they have computed the distance of a point on an object from camera and have shown that the computed distance from camera is almost equal to the actual distance from camera.

Lee et al. (Lee et al., 1994) used perspective transformation procedure for mapping a 3-D scene onto an image plane. It has also been shown that the missing depth information can be obtained by using stereoscopic imaging techniques. They derived an equation for finding the depth information using 3-D geometry. The most difficult task in using the equation for

obtaining the depth information is to actually find two corresponding points in different images of the same scene. Since, these points are generally in the same vicinity, a frequently used approach is to select a point within a small region in one of the image views and then attempt to find the best matching region in the other view by using correlation techniques. The geometry of a 3-D scene can be found if it is known which point from one image corresponds to a point in the second image. This correspondence problem can be reduced by using several constraints.

Klette et al. (Klette et al., 1996) proposed a list of constraints like uniqueness constraint, photometric compatibility constraint, and geometric similarity constraint etc that are commonly used to provide insight into the correspondence problem. They illustrated this approach with a simple algorithm called block matching. The basic idea of this algorithm is that all pixels in the window (called a block) have the same disparity, meaning that one and only one disparity is computed for each block.

But later on Nishihara (Nishihara, 1984) noted that such point-wise correlators are very heavy on processing time in arriving at a correspondence. He proposed another relevant approach to match large patches at a large scale, and then refine the quality of the match by reducing the scale using the coarser information to initialize the finer-grained match.

Pollard et al. (Pollard et al., 1981) developed the PMF algorithm using the feature-based correspondence method. This method uses points or set of points those are striking and easy to find, such as pixels on edges, lines or corners. They proceed by assuming that a set of feature points [detected edges] has been extracted from each image by some internal operator. The output is a correspondence between pairs of such points.

Tarabanis et al. (Tarabanis & Tsai, 1991) described the next view planning method as follows: "Given the information about the environment as well as the information about that the vision system has to accomplish (i.e. detection of certain object features, object recognition, scene reconstruction, object manipulation), develop strategies to automatically determine sensor parameter values that achieve this task with a certain degree of satisfaction".

Maver et al. (Maver & Bajcsy, 1993) proposed an NVP (Next View Planning) algorithm for an acquisition system consisting of a light stripe range scanner and a turntable. They represent the unseen portions of the viewing volume as 2½-D polygons. The polygon boundaries are used to determine the visibility of unseen portions from all the next views. The view, which can see the largest area unseen up to that point, is selected as the next best view.

Connolly (Connolly, 1985) used an octree to represent the viewing volume. An octree node close to the scanned surface was labeled to be seen, a node between the sensor and this surface as empty and the remaining nodes as unseen. Next best view was chosen from a sphere surrounding the object.

Szeliski (Szeliski, 1993) first created a low-resolution octree model quickly and then refined this model iteratively, by intersecting each new silhouette with the already existing model. Niem (Niem, 1994) uses pillar-like volume elements instead of an octree for the model representation.

Whaite et al. (Whaite & Ferrie, 1994) used the range data sensed to build a parametric approximate model of the object. But this approach does not check for occlusions and does not work well with complex objects because of limitations of a parametric model.

Pito (Pito, 1999) used a range scanner, which moves on a cylindrical path around the object. The next best view is chosen as the position of the scanner, which samples as many void patches as possible while resampling at least a certain amount of the current model.

Liska (Liska, 1999) used a system consisting of two lasers projecting a plane onto the viewing volume and a turntable. The next position of the turntable is computed based on information from the current and the preceding scan.

Sablating et al. (Sablating et al., 2003; Lacquaniti & Caminiti, 1998). described the basic shape from Silhouette method used to perform the 3-D model reconstruction. They experimented with both synthetic and real data.

Lacquaniti et al. (Lacquaniti & Caminiti, 1998) reviewed anatomical and neurophysical data processing of a human in eye-memory during grasping. They also established the different mapping techniques for ocular and arm co-ordination in a common reference plane.

Desai (Desai, 1998) in his thesis addressed the problem of motion planning for cooperative robotic systems. They solved the dynamic motion-planning problem for a system of cooperating robots in the presence of geometric and kinematic constraints with the aid of eye memory co ordination.

Metta et al. (Metta & Fitzpatrick, 2002) highlighted the sensory representations used by the brain during reaching, grasping and object recognition. According to them a robot can familiarize itself with the objects in its environment by acting upon them. They developed an environment that allows for a very natural developmental of visual competence for eye-memory prehension.

Barnesand et al. (Barnesand & Liu, 2004) developed a philosophical and psychophysiological basis for embodied perception and a framework for conceptual embodiment of vision-guided robots. They argued that categorization is important in all stages of robot vision. Further, classical computer vision is not suitable for this categorization; however, through conceptual embodiment active perception can be erected.

Kragic et al. (Kragic & Christensen, 2003) considered typical manipulation tasks in terms of a service robot framework. Given a task at hand, such as "pick up the cup from the dinner table", they presented a number of different visual systems required to accomplish the task. A standard robot platform with a PUMA560 on the top is used for experimental evaluation.

The classical approach-align-grasp idea was used to design a manipulation system (Bhaumik et al., 2003). Both visual and tactile feedback was used to accomplish the given task. In terms of image processing, they started by a recognition system, which provides a 2-D estimate of the object position in the image. Thereafter, a 2-D tracking system was presented and used to maintain the object in the field of view during an approach stage. For the alignment stage, two systems are available. The first is a model based tracking system that estimates the complete pose/velocity of the object. The second system was based on corner matching and estimates homography (matching of periphery) between two images. In terms of tactile feedback, they presented a grasping system that performs power grasps. The main objective was to compensate for minor errors in object's position/orientation estimate caused by the vision system.

Nakabo et al. (Nakabo et al., 2002) considered real-world applications of robot control with visual servoing; both 3-D information and a high feedback rate is required. They developed a 3-D target tracking system with two high-speed vision systems called Column Parallel Vision (CPV) systems. To obtain 3-D information, such as position, orientation and shape parameters of the target object, a feature-based algorithm has been introduced using moment feature values extracted from vision systems for a spherical object model.

### 3. Objective

In the present investigation, an attempt has been made to enable a four-fingered robotic hand consisting of the index finger, middle finger, ring finger and the thumb to ensure stable grasp. The coordinated movement of the fingertips were thoroughly analyzed to preshape the fingers during trajectory planning in order to reduce task execution time. Since the displacement of the motor was coordinated with the motion of the fingertips, the correlation between these two parameters was readily available through CAD simulation using Visual Nastran 4D (MSC. VisualNastrun 4D, 2001).

The primary objectives of the present investigation are:

- a) analysis of the object shapes and dimensions using 2D image processing techniques and vision based preshaping of the finger's pose depending on the basis of prehension,
- b) hierarchical control strategies under vision guidance for slip feedback,
- c) experimentation on the hand for intelligent grasping,

### 4. Brief Description of the Setup for Prehension

#### 4.1 Kinematics of the Hand

The newly developed hand uses a direct linkage mechanism to transfer motions. From Fig.4 it is clear that the crank is given the prime motion, which ultimately presses the push link-1 to turn the middle link about the proximal joint. As the middle link starts rotating, it turns the distal link simultaneously, because the distal link, middle link and push link-2 form a crossed four bar linkage mechanism. The simultaneous movement of the links ultimately

stops when the lock pin comes in contact with the proximal link. The proximal link is kept in contact with the stay under the action of a torsional spring. As the lock pin comes in contact with the proximal link, it ultimately restricts all the relative motions between the links of the finger and at that position the whole finger moves as an integrated part. The crank is coupled to a small axle on which a worm wheel is mounted and the worm is directly coupled to a motor by coupling as shown in Fig.5.

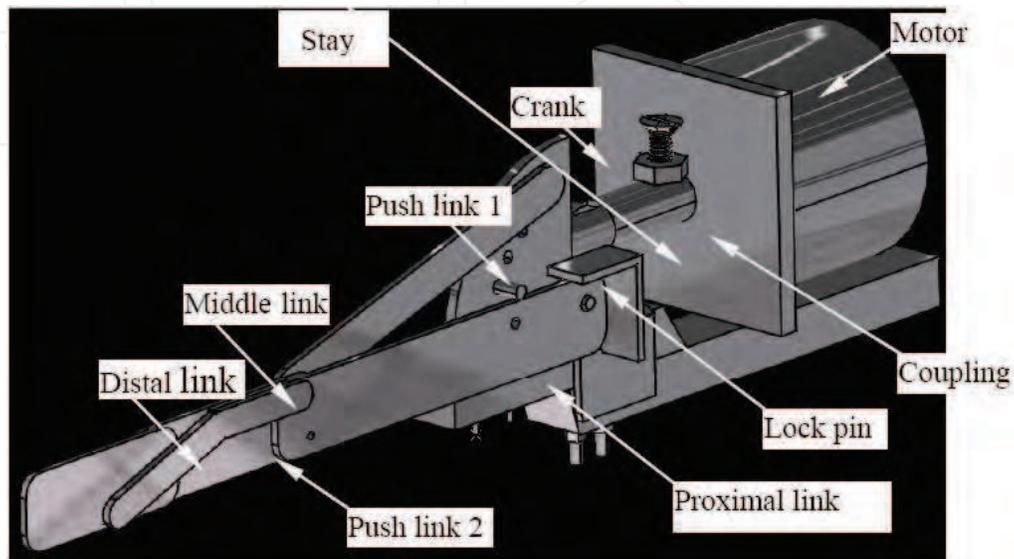


Fig. 4. The kinematic linkages for the robot hand

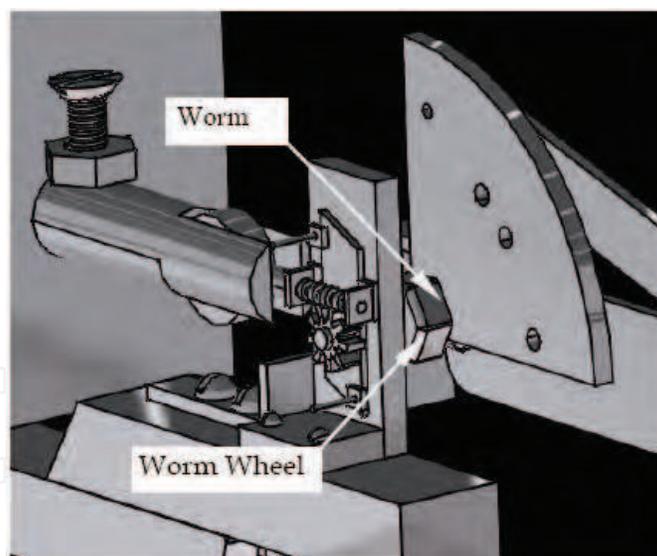


Fig. 5. The worm worm-wheel mechanism for actuation of the fingers

#### 4.2 Description of the Mechanical System

The robotic hand consists of four fingers namely thumb, index, middle and ring fingers. Besides, there is a palm to accommodate the fingers along with the individual drive systems for actuation. The base, column and the swiveling arm were constructed in this system for supporting the hand to perform the experiments on grasping. The design of the base, column and swiveling arm were evolved from the absence of a suitable robot.

The base is mild steel flat on which the column is mounted. On the free end of the column, a provision (hinge) has been made to accommodate the swiveling arm adaptor. The CAD model of the setup has been shown in Fig.6.

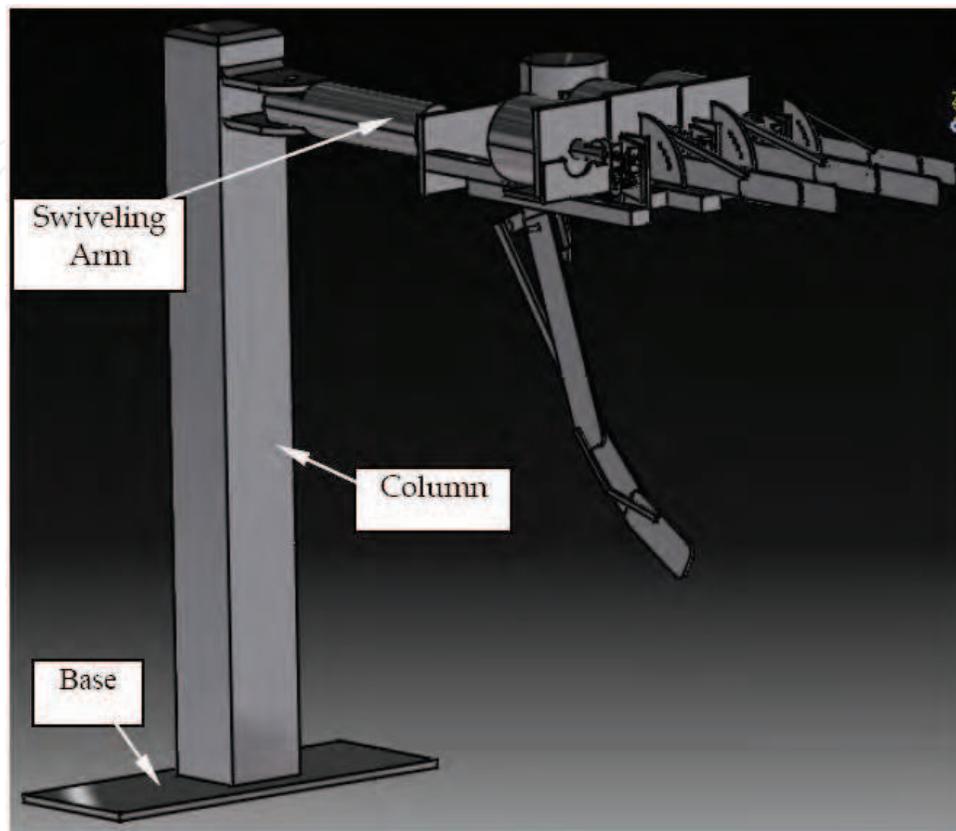


Fig. 6. The assembled CAD Model of the robot hand

### 4.3 Vision System

The specification of the camera is as follows:

- |                              |   |                             |
|------------------------------|---|-----------------------------|
| a. Effective picture element | : | 752 (H) × 582 (V)           |
| b. Horizontal frequency      | : | 15.625 KHz ± 1%             |
| c. Vertical Frequency        | : | 50Hz ± 1%                   |
| d. Power requirements        | : | 12v DC ± 1%                 |
| e. Dimension                 | : | 44(W) × 29 (H) × 57.5(D) mm |
| f. Weight                    | : | 110 gm                      |

A two-channel PC based image-grabbing card was used to acquire the image data through the camera. Sherlock™ (Sherlock) is a window based machine vision environment specifically intended to simplify development and deployment of high performance alignment, gauging inspection, assembly verification, and machine guidance tasks. This was used to detect all the peripheral pixels of the object being grasped after thresholding and calibration. The dimensional attributes are solely dependant on the calibration. After calibration a database was made for all the peripheral pixels. The camera along with the mounting device has been shown in Fig.7.



Fig. 7. The Camera alongwith the mounting device

### 5. Fingertip Trajectories and Acquisition of Pre-shape Values

To determine the trajectories of the fingers in space, points on the distal palmer tip were chosen for each of the fingers and during simulation the locus of those points were traced. The instantaneous crank positions were also taken simultaneously. As the fingers flex, the coordinate abscissa ( $X$ ) value either increases or decreases depending on the position of the origin as the current system. Since the incremental values for coordinate ( $X$ ) are correlated with the angular movement of the crank, during preshape of the fingers, proper actuation can be made as shown in Fig.8. Once the vision based sensory feedback values are known, the motors may be actuated to perform the required amount of incremental movements.

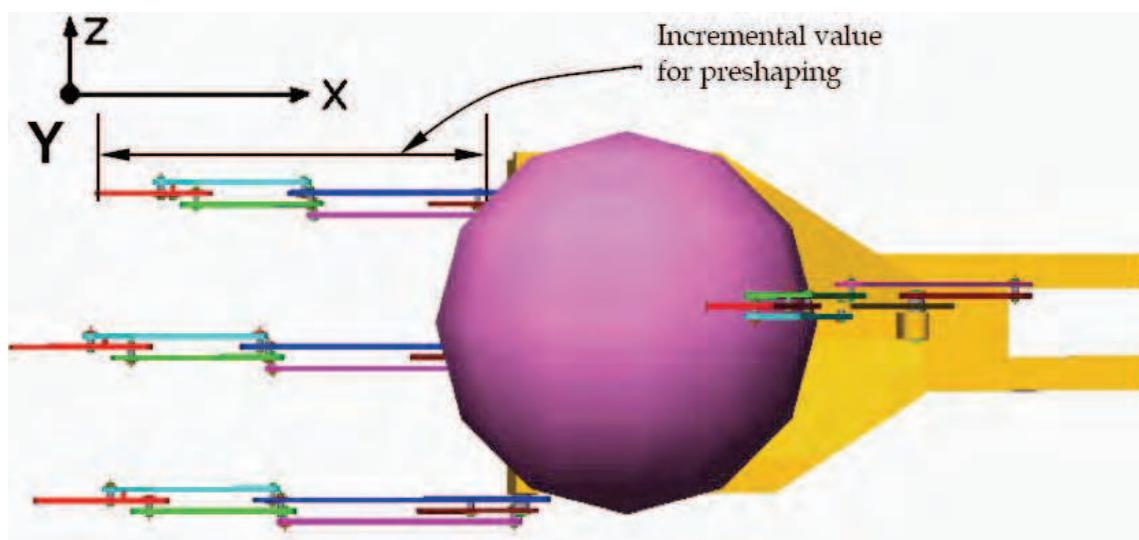


Fig. 8. The preshape value for the fingers

The model was so made that the direction of the finger axis was the  $X$ -axis and the gravity direction was the  $Y$  direction. Figure 9 shows the trajectories for different fingers. The correlations of the incremental values in preshaping direction and the corresponding crank movement have been shown in Fig.10, Fig.11, Fig.12 and Fig.13. The  $R^2$  values in the curves imply the correlation constant and as the value tends to 1 (one) implies a good correlation of the curve fitting.

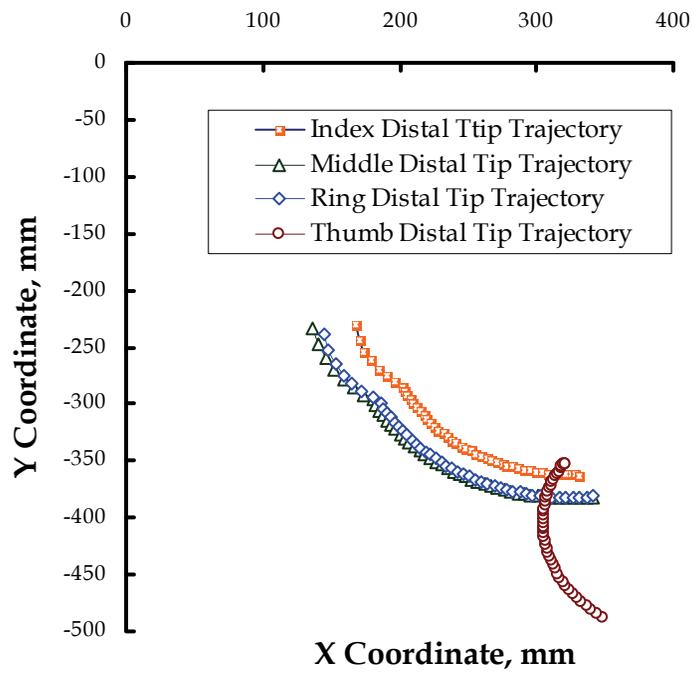


Fig. 9. The trajectories of the fingertips of the fingers

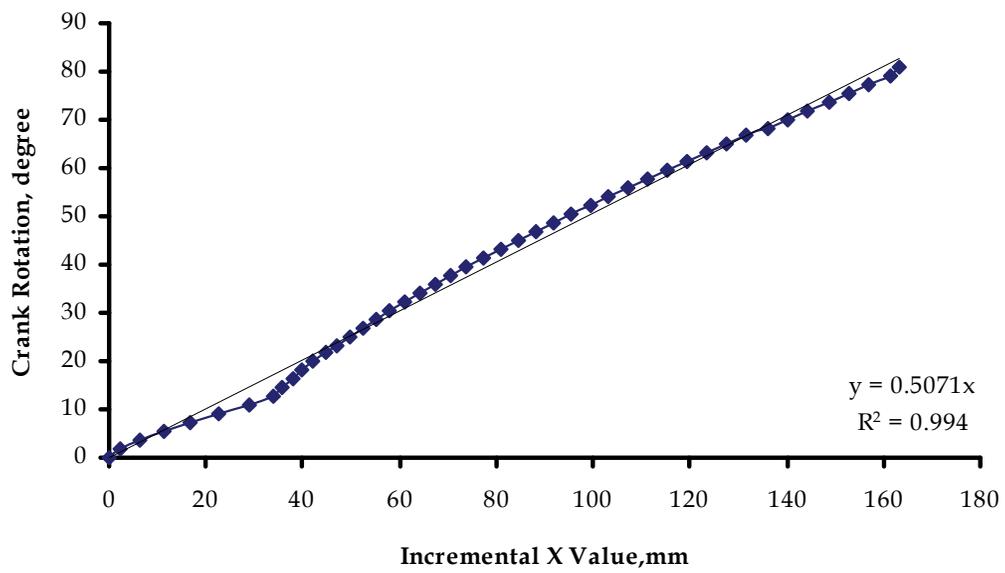


Fig. 10. Preshape value vs. crank rotation angle plot for index finger

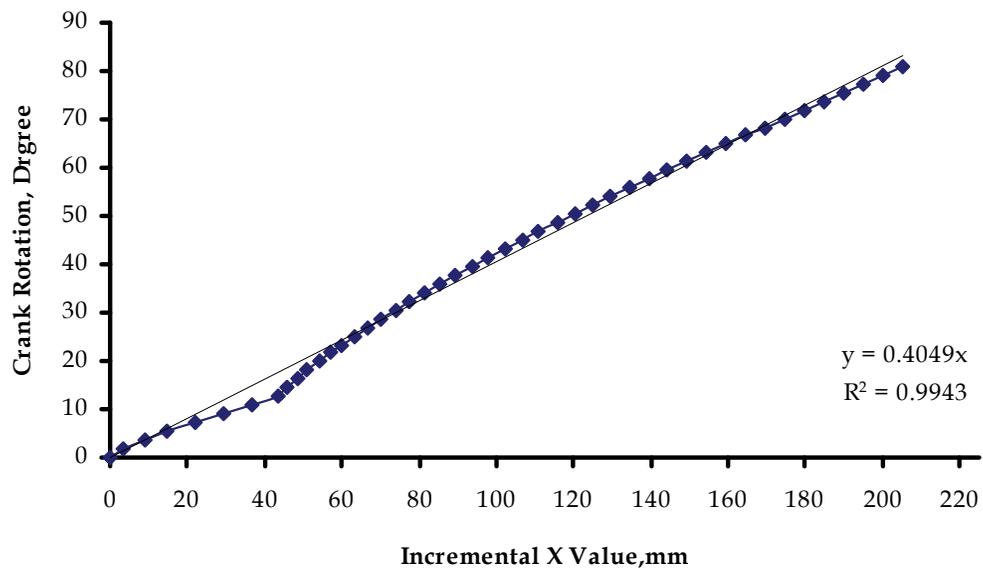


Fig. 11. Preshape value vs. crank rotation angle plot for middle finger

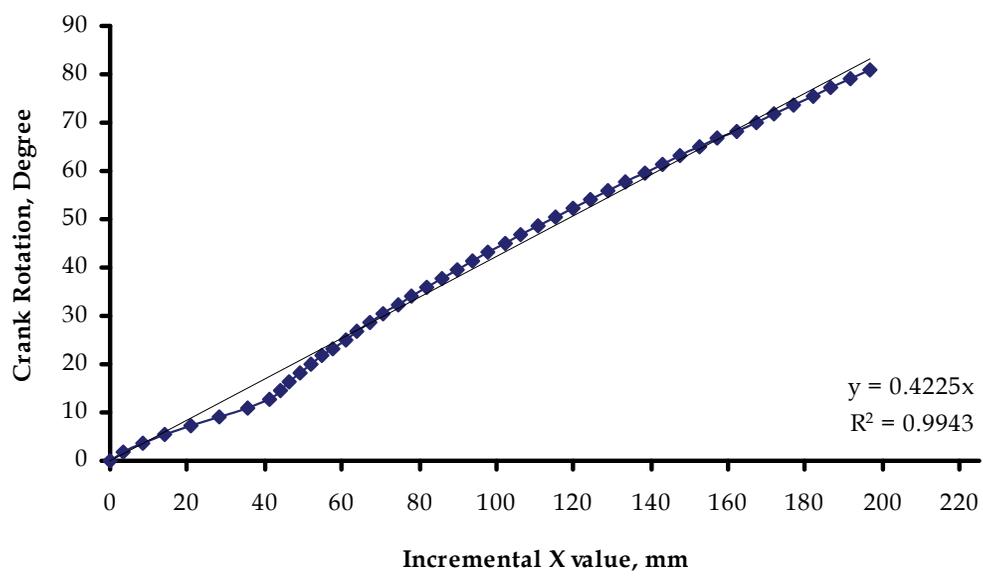


Fig. 12. Preshape value vs. crank rotation angle plot for ring finger

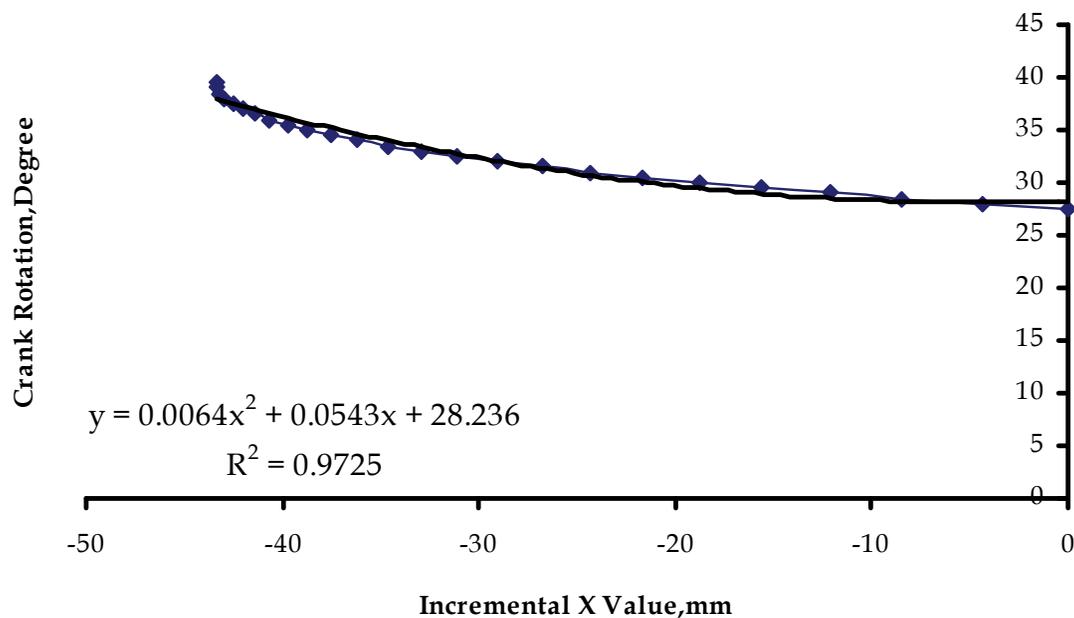


Fig. 13. Plot showing preshape value vs. crank rotation angle for thumb

## 6. Control Hierarchy and Vision Assistance

The control architecture is a three level hierarchy, which consists of the following blocks:

### 6.1 A computer with CCD Camera i.e. Master Computer-I

The computer at the top level of hierarchy gives the object shape information to the next level computer to actuate the motors for preshaping through the slave i.e. the microcontroller motherboard.

### 6.2 The Master Computer-II

The second level computer acting in master mode communicates with the micro-controller motherboard that acts as slave through the serial port, RS 232.

### 6.3 Controller Mother Board with Micro Controller Chip (AT89C52)

The main function of micro-controller motherboard is to behave like that of a Data Acquisition System (DAS) and thus helps the master by providing the required data from different sensors. Evaluating the controller mode equation through software, the master downloads the firing angle information and motor direction information in to the micro-controller chip (AT89C52), which in turn sends these signals to the drive circuitry. Figure 14 represents the realistic view for the three-generation controller. Figure 15 depicts the overall control architecture used in this scheme.

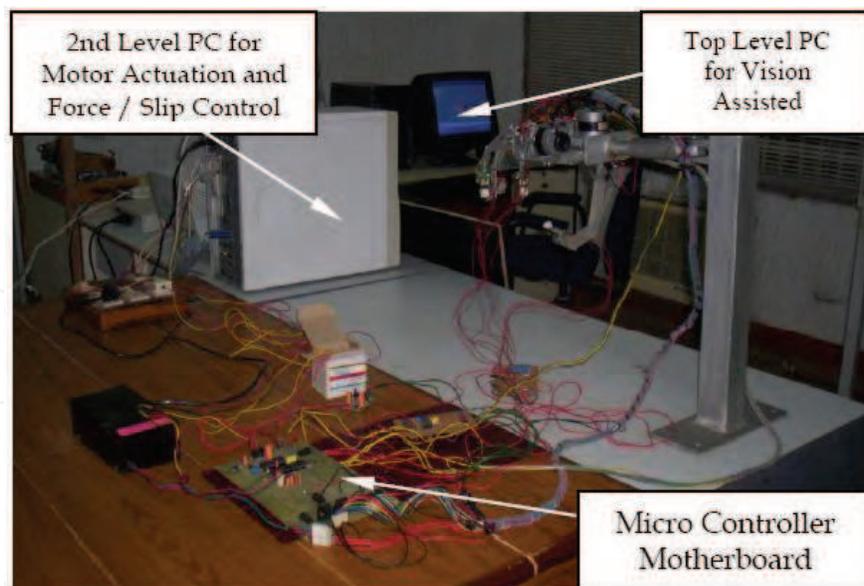


Fig. 14. Prehension system with three level hierarchy

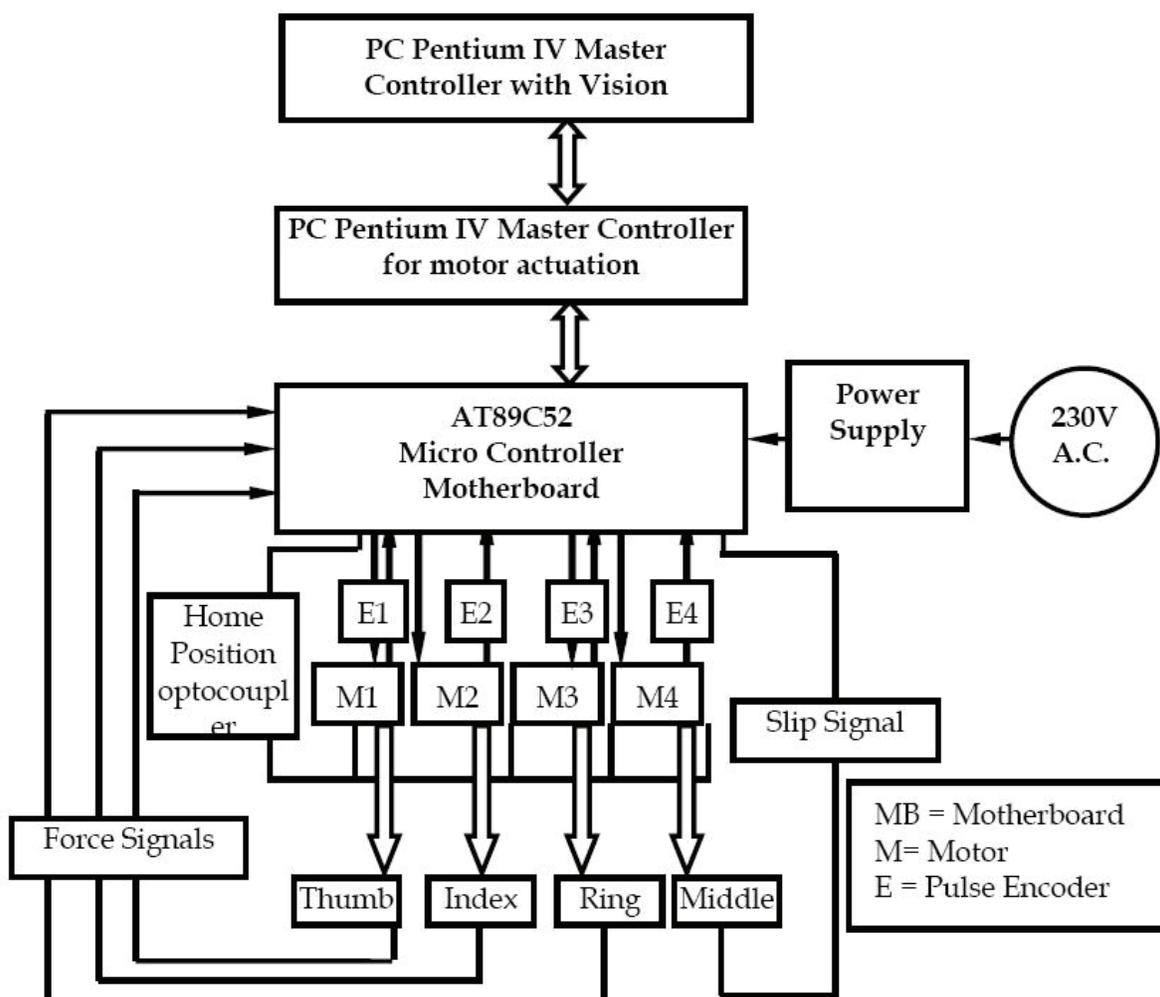


Fig. 15. The three level hierarchy of control system

#### 6.4 Actuation of Motors under Vision Assistance

The motors are acting as actuators for finger motion as the finger positions are determined by the motor position. The vision assistance helps the fingers to preshape to a particular distance so that the fingertips are a few units apart from the object. As soon as the fingers touch the object, the force sensor reads the current value of the force exerted by all the fingers. The incremental step is given to all the motors as  $7.2^\circ$  i.e. four counts.

The control strategy is divided into three steps:

- The object shape information can be had by CCD camera to generate the amount of preshape (i.e. the pulse count by which the motor will rotate to encompass the job). The initial set value for the position comes from the vision assistance. The scheme has been depicted in Fig.16.
- Secondly it is to generate a database for grasping force exerted by individual finger for grasping different materials under no slip condition. The amount of slip is then fed back to the computer and the computer through microcontroller sends signal to the motors via motor drive card to move forward. The total control strategy to generate the database of force at no slip condition has been shown in Fig.16.

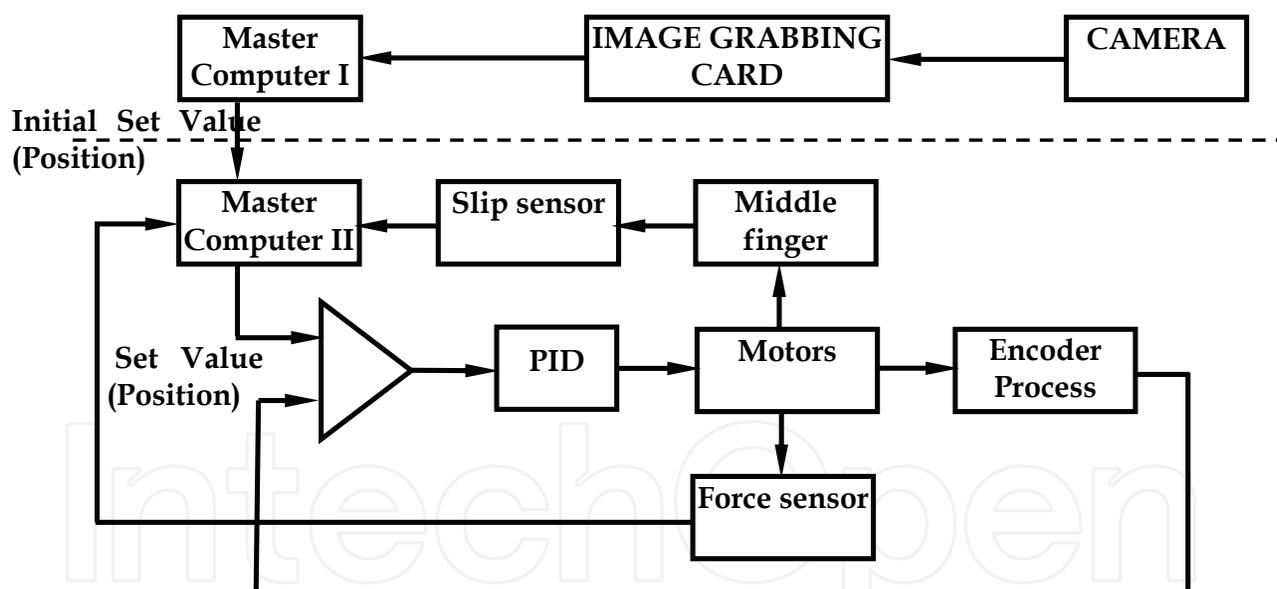


Fig. 16. The vision based preshaping and slip feedback

Then the values of the grasping forces at which slip stops is feed to the controller i.e. the grasping force exerted by the individual fingers for different objects. Once the set value is reached then the controller simultaneously checks whether there is any slip or not. If till now the fingers are not capable to grip, then the set values are dynamically increased to a maximum of 10% of the initial set values with an increment of 1%. Beyond which the system will declare that the process is out of range.

This process approaches towards adaptive control. The schematic diagram of force control has been shown in Fig.17. The microcontroller motherboard is connected to the computer by RS-232 serial port. But position forward scheme does not imply that there will always be a force signal, rather the fingertips are to be in contact with the object at first, only then the force signals can be achieved.

In order to control the grasp position of the gripper for holding an object securely the motor should achieve the desired set position smoothly without any overshoot and undershoot obviously with negligible offset error. This set position corresponds to the stable grasp position. To control the gripping force required for an object, a PID closed loop feedback control scheme is adopted. The initial set value for position comes from the vision data and the motors rotate to preshape the objects dimension and to see whether the gripper is in contact with the object or not. If the gripper is not in contact with the object then the motor moves forward till all the fingers come in contact with the job. If the contact exists, then the PID controller sets the current process value plus four counts as set value and checks whether there is any slip or not. Now if there is a slip, the set value is increased by another four counts until no slip condition is reached and the master computer-II saves all the optimum grasping value in the database. The flow chart has been depicted in Fig.18.

In the second scheme i.e. the force control scheme, the initial set values for force comes from the database of the master computer. Feed forward signal are given to the motors for wrapping the object. This process continues till the desired set values for all the fingers are reached. Then it is ensured whether there is any slip or not after reaching to the set position. If till now there is any slip then the set value will be set as process value plus 1% of the set value until slippage stops and gripper holds the object securely. Each fingertip contains electrical strain gauge type force sensing elements for the measurement of the forces. The control flowchart scheme has been shown in Fig.19.

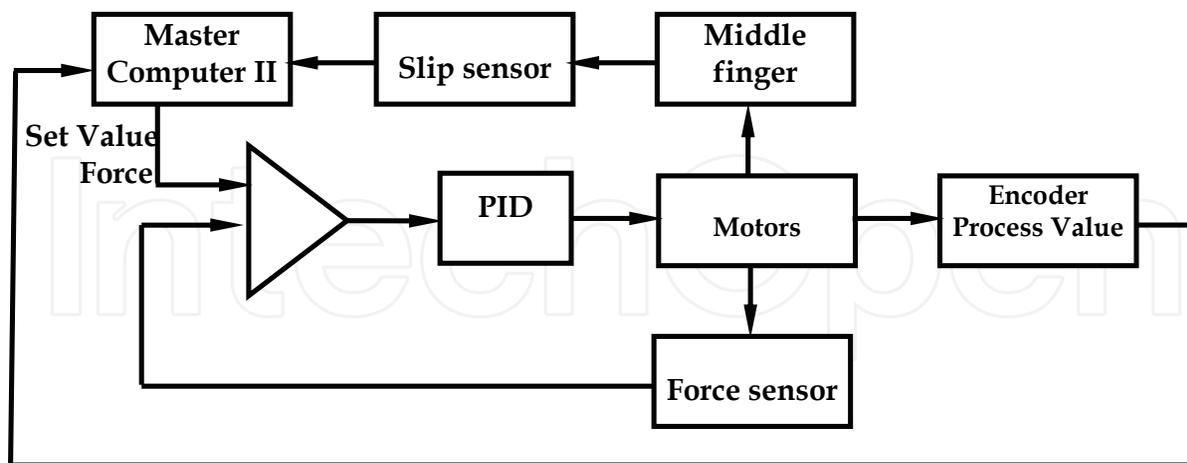


Fig. 17. The force feedback for stable grasp

The main advantage of this composite control, PID mode is that the integral mode eliminates the offset error produced by proportional mode and the overshoot and undershoot of the controlled output is minimized by derivative action through its anticipatory property. The overall effect is that the output-controlled variable achieves the

desired value without much cycling. Critically damping condition has been implemented in the present work. The control software has got the feature to set the set value manually or automatically (from the database). The gain values can also be set from the software as shown in Fig.20.

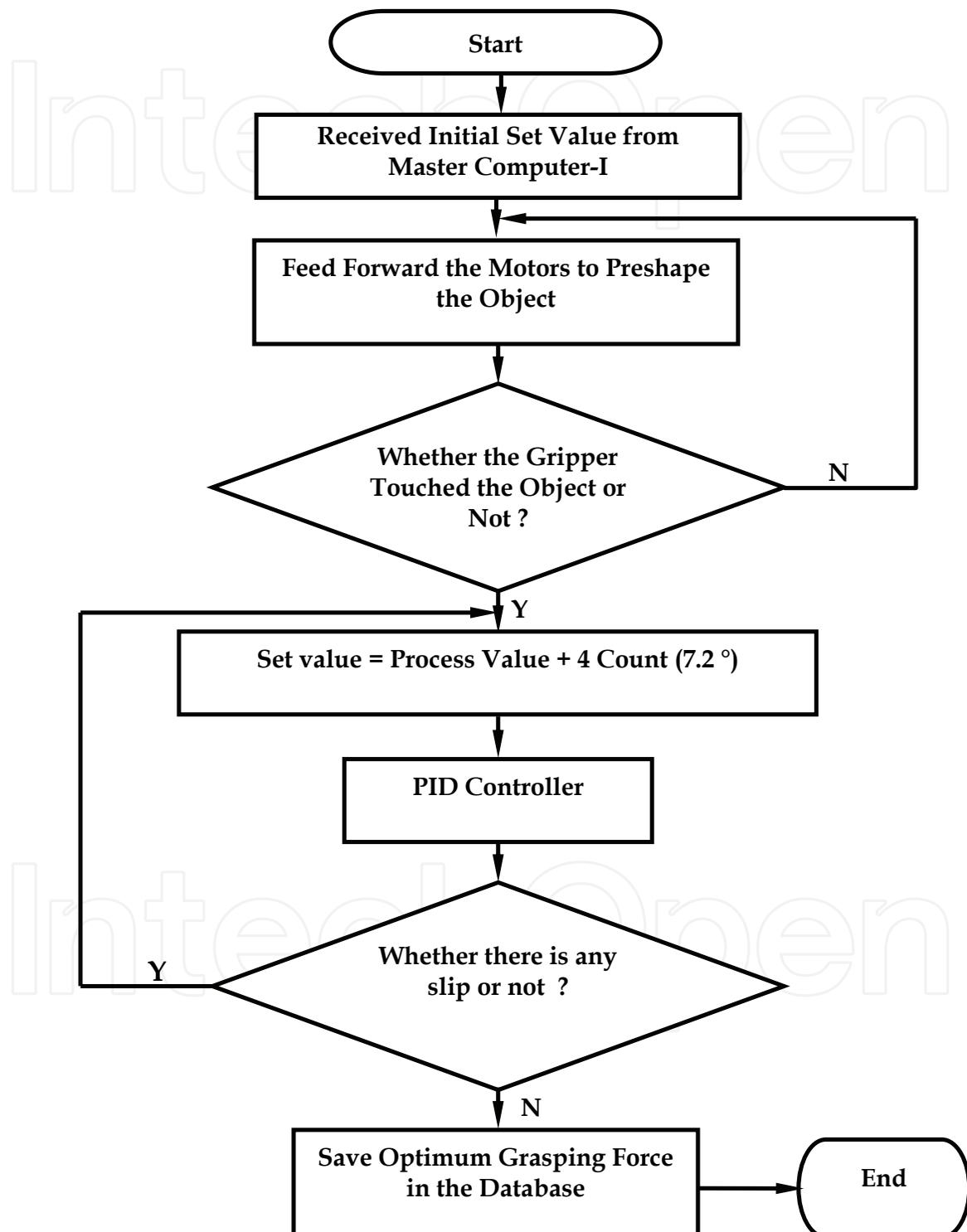


Fig. 18. The flowchart depicting vision assisted preshape and slip feedback

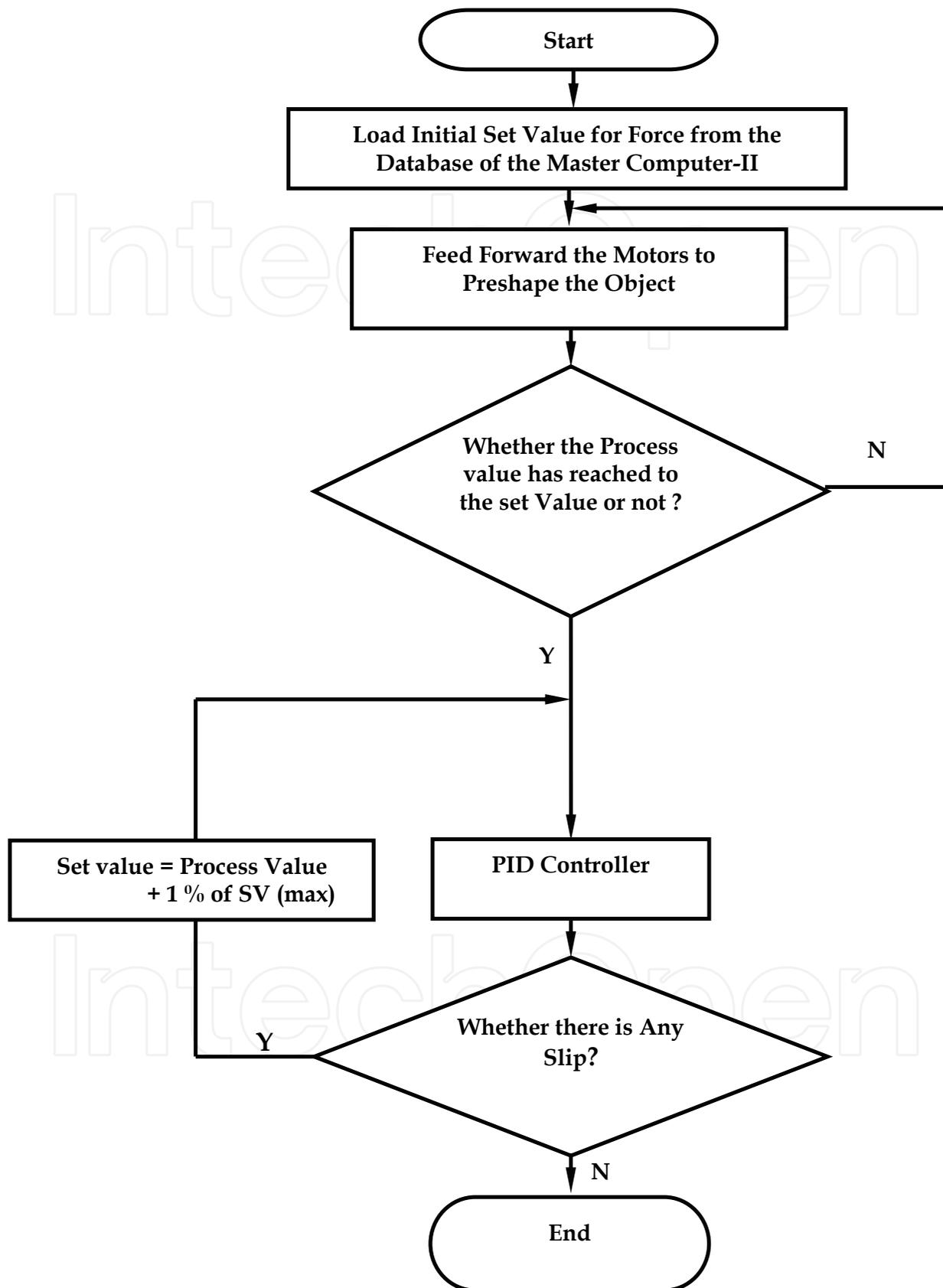


Fig. 19. The flowchart showing force feedback

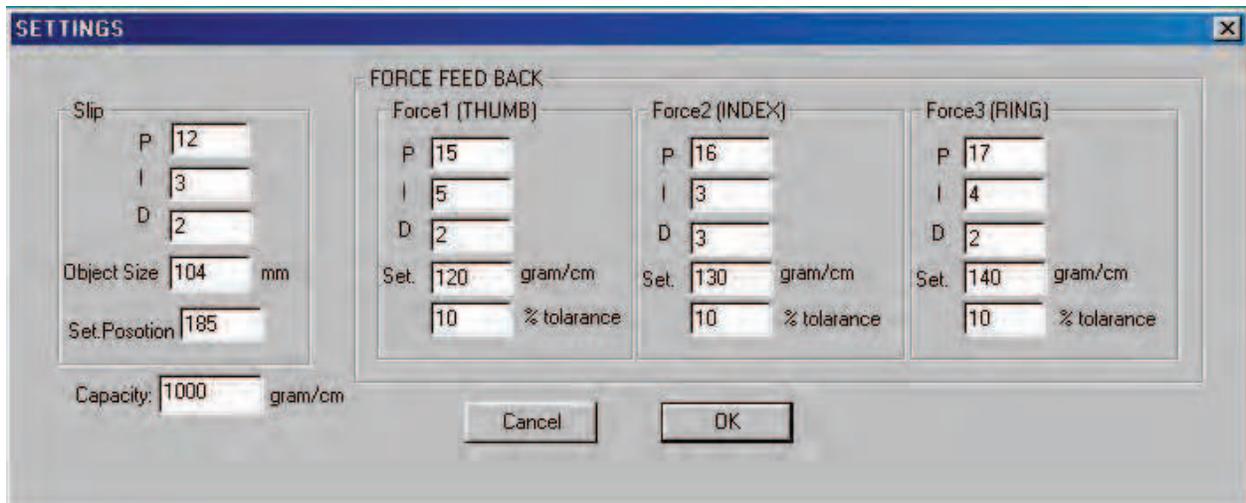


Fig. 20. Set values for both slip and force

A CCD camera has been used to generate the information about the object shape. The home position of the fingers before grasping is known. A logic has been developed to find out the movements of the corresponding motors of the fingers in terms of P1 and P2 as shown in Fig 21.

In article 5 the correlation between the motor position and incremental preshape value has already been found out from the correlation equations indicated in Figs 10, 11, 12 and 13. The Master Computer-II acquires the dimension from the Master Computer-I and calculates the preshape. In the software a provision has been made to input either the object size data obtained from the visual information incorporate user input data for the geometry of the primitive encompassing the object. The data for object size has been illustrated in Fig.22.

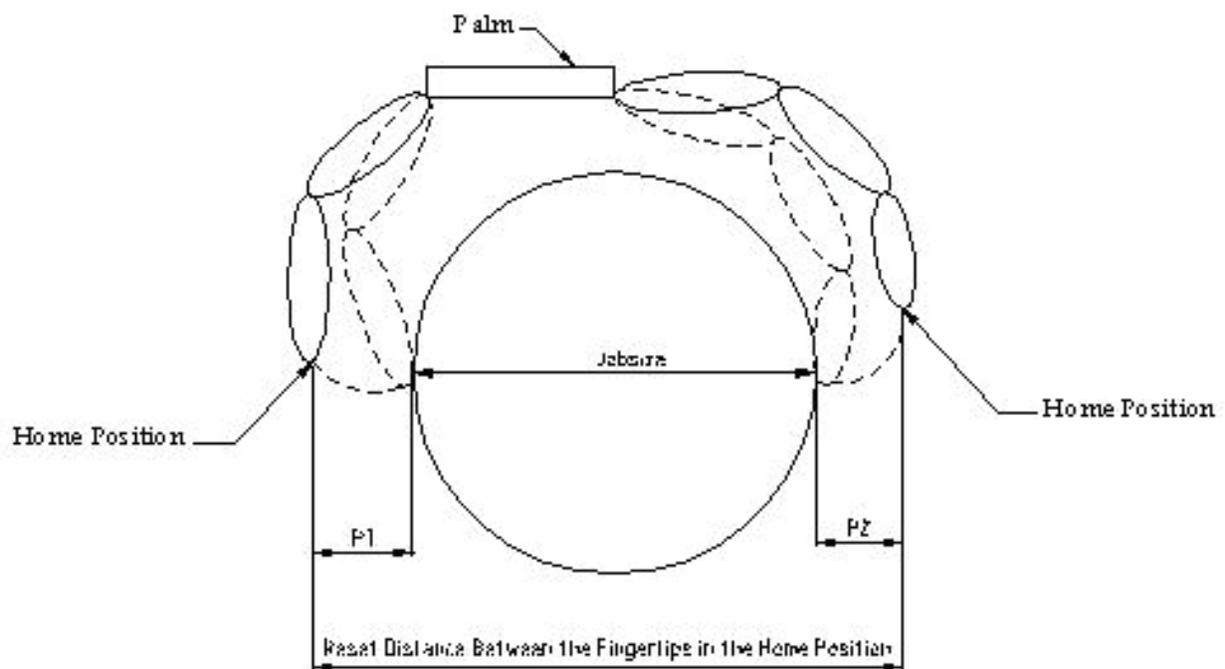


Fig. 21. The preshape calculation for the fingers

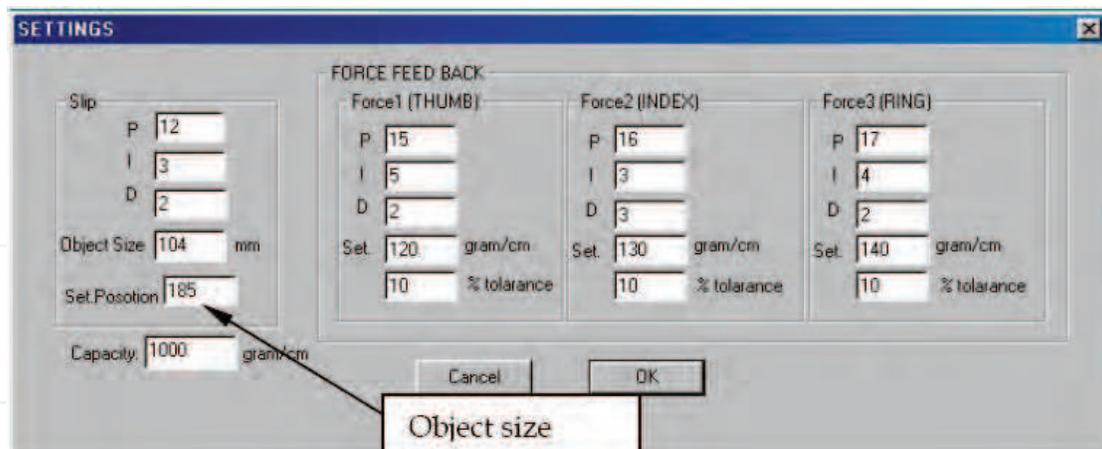


Fig. 22. The provision for incorporation of the object size

## 7. Grasp Stability Analysis through Vision

### 7.1 Logic for Stable Grasp

When an object is held through a robot hand, to maintain stability it becomes very important that the line of action of the force exerted by the finger tip should lie within the cone of friction shown in Fig.23. As it is well known that the cone of friction is again a function of the coefficient of friction. For this reason it has been checked that, whether the line of action of the forces are lying with the cone of friction or not.

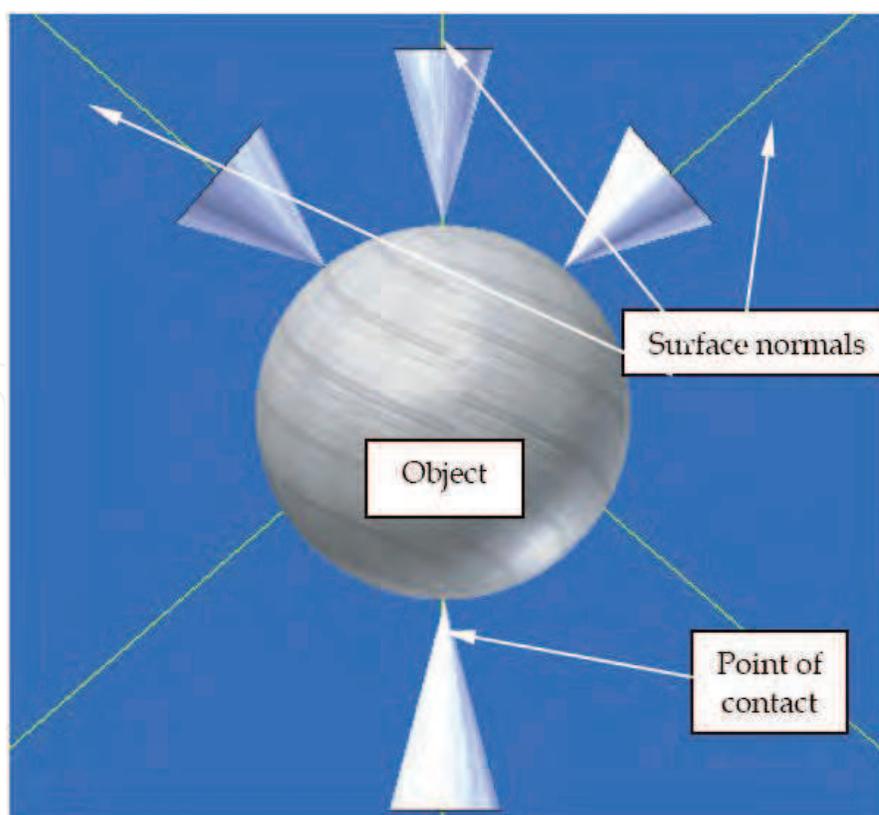


Fig. 23. The relative position of the fingertips with respect to the object

## 7.2 Interpolation of Cubic Spline to form the mathematical Silhouette

In order to find out instantaneous normal on the silhouette a piece-wise cubic spline was interpolated amongst the peripheral points. The logic for cubic spline is as follows:

For a set of data (had from silhouette) points it is necessary to inculcate the mode of connectivity and order of joining of the points-this in turn yield the idea about continuity when a piecewise polynomial curve come in the design scenario. By continuity we mean the index of smoothness at the juncture of two pieces of a curve. Theoretically a continuous curve means, within the boundary of the independent variable the dependent variables are fully predictive. At the same time polynomial of degree N, has a continuity of derivatives of order (N-1). Parametric cubic splines are used to interpolate two given data, not to design free form curves as Bezier and  $\beta$ -Spline curves do. The Hermite form of Cubic Spline is determined by defining positions and tangent vectors at the data points. Therefore, four conditions are required to determine the co-efficients of the parametric Cubic Spline curve. When these are the positions of the two end points and the two tangent vectors at the points, a hermite cubic spline results.

The parametric equation of the Cubic Spline segment is given by,

$$P(u) = \sum_{i=1}^3 C_i u^i, \quad 0 \leq u \leq 1 \quad (1)$$

Where  $u$  is the parameter and  $C_i$  is the polynomial co-efficient.

The equation 1 can be written in an expanded form as

$$P(u) = C_3 u^3 + C_2 u^2 + C_1 u + C_0 \quad (2)$$

This equation can also be written in a matrix form as

$$P(u) = U^T C \quad (3)$$

Where,  $U = [u^3 \quad u^2 \quad u \quad 1]^T$  and  $C = [C_3 \quad C_2 \quad C_1 \quad C_0]^T$

The tangent vector to the curve at any point is given by

$$P'(u) = \sum_{i=0}^3 C_i i u^{i-1} \quad 0 \leq u \leq 1 \quad (4)$$

The co-efficients are to be replaced by the geometry of the spline. In order to eliminate the co-efficients two end points  $P(0)$  and  $P(1)$  (as shown in the adjoining Fig.24) and the two end tangents  $P'(0)$  and  $P'(1)$  are to be given .

$$P(0) = C_0, \quad (5)$$

$$P'(0) = C_1 \quad (6)$$

$$P(1) = C_3 + C_2 + C_1 + C_0 \quad (7)$$

$$P'(1) = 3 C_3 + 2 C_2 + C_1 \quad (8)$$

Where,  $P(0)$  = Coordinates of the first point at  $t = 0$

$P(1)$  = Coordinates of the last point at  $t = 1$

$P'(0)$  = Values of the slopes in  $x, y, z$  directions at  $t = 0$

$P'(1)$  = Values of the slopes in  $x, y, z$  directions at  $t = 1$

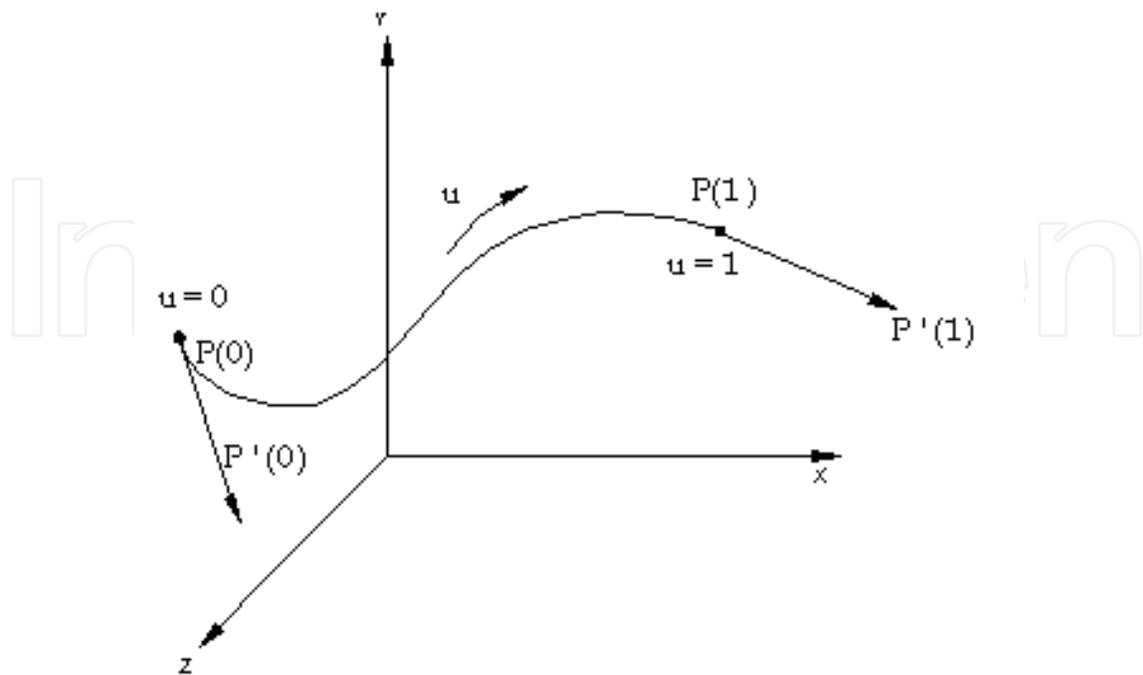


Fig. 24. Parametric representation of cubic spline

The solution of the four simultaneous equations written above for the coefficients, results the following equations

$$C_0 = P(0) \quad (9)$$

$$C_1 = P'(0) \quad (10)$$

$$C_2 = 3[P(1) - P(0)] - 2[P'(0) - P'(1)] \quad (11)$$

$$C_3 = 2[P(0) - P(1)] + P'(0) + P'(1) \quad (12)$$

Now substituting these values of the coefficients into the equation 2, we get

$$P(u) = (2u^3 - 3u^2 + 1)P_0 + (-2u^3 + 3u^2)P_1 + (u^3 - 2u^2 + u)P'_0 + (u^3 - u^2)P'_1, \quad 0 \leq u \leq 1 \quad (13)$$

Differentiating equation 13 the tangent vector can be obtained,

$$P'(u) = (6u^2 - 6u)P_0 + (-6u^2 + 6u)P_1 + (3u^2 - 4u + 1)P'_0 + (3u^2 - 2u)P'_1 \quad 0 \leq u \leq 1 \quad (14)$$

The functions of  $u$  in the equation 13 and 14 are called blending functions.

Equation 13 can be written in the matrix form as,

$$P(u) = U^T[M_H]V, \quad 0 \leq u \leq 1 \quad (15)$$

Where,  $[M_H]$  is the Hermite matrix and  $V$  is the geometry (or boundary conditions) vectors. Both are given by

$$[M_H] = \begin{bmatrix} 2 & -2 & 1 & 1 \\ -3 & 3 & -2 & -1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \quad (16)$$

$$V = [P_0 \quad P_1 \quad P'_0 \quad P'_1]^T \quad (17)$$

Comparing eq<sup>n</sup> 3 and eq<sup>n</sup> 15, we can see that,

$$C = [M_H]V \text{ or } V = [M_H]^{-1}C \quad (18)$$

Now eq<sup>n</sup> 13 shows that a cubic spline curve can be represented by its two endpoints and two tangent vectors. The curve passes through the end points ( $u=0$  and  $1$ ) and from these above equations; it can be shown that, changing the end points or the tangent vectors, the shape of the curve can be controlled.

### 7.3 Algorithm for Stable Grasp

A suitable algorithm has been developed for thresholding and edge detection (Bepari, 2006; Datta & Ray, 2007). After finding out the silhouette, a virtual hand is utilized for grasping the periphery of the object. At every instance the object rotates by an angle of  $1^\circ$  and the angle between line of action of the force exerted by the finger and the instantaneous surface normal at that point of contacts are stored for optimization purpose. If this angle is less than that of the cone of friction of the work-hand pair, then the mapping is accepted otherwise rejected.

The said procedure iterates until the optimized grasp is obtained. The object snap, threshold figure and the optimal grasp have been shown in Fig.25. The contact property can be changed by changing the co-efficient of friction. The software has been developed by using VISUAL BASIC. Another constraint during design of stable grasp is the maximum distance between object's centre of gravity and hand's centre of gravity shown in the same figure and the absolute values of the coordinates of the two said points may be selected for stable orientation.

The methodology for stable grasp is as follows:

- i. Snap of the object (in the plane of grasp or from the direction normal to the grasping plane),
- ii. Generate the binary image,
- iii. Calibrate the image,

- iv. Generate edge detected boundary points using band-thresholding,
- v. Cubic spline polynomials are interpolated,
- vi. Input user input data (the translation zone of the end effector i.e. the starting coordinate and end coordinate enclosing a rectangle, coefficient of friction,  $\mu$ , step angle, translational step),
- vii. Specify the line of action of the force of the fingertips i.e. the plane of finger actuation,

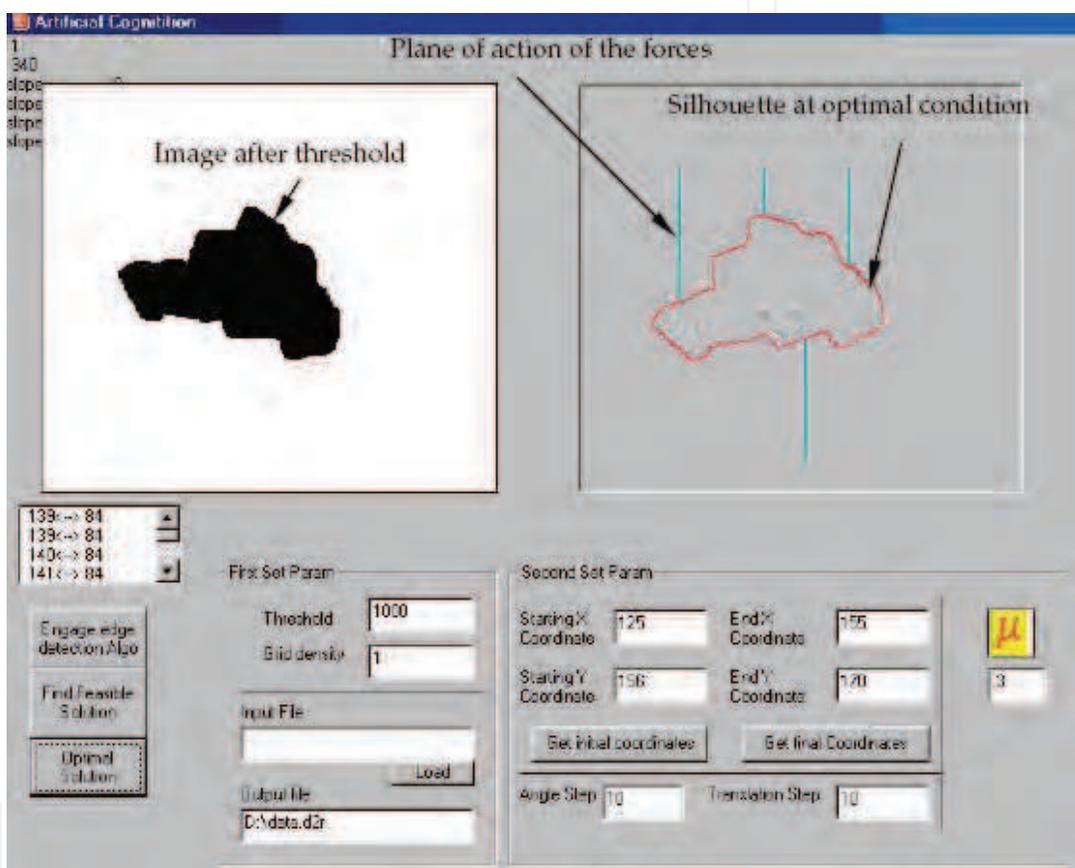


Fig. 25. The optimized grasp orientation

- viii. Generate the feasible solution,
  - a. Simultaneously rotate and translate the boundary points,
  - b. At each step of rotation the entire finger plane intersects at points with the profile of the object are found,
  - c. Corresponding surface normal is found,
  - d. The angles between the normal at those points and the line of action of the force are calculated.

- e. Since  $\mu$  is given it is checked whether the angle of the cone of friction is less than  $\tan^{-1}(\mu)$  or not,
  - f. If it is less than  $\tan^{-1}(\mu)$  for all the fingers then angle of each finger and the coordinates of the figure are stored in a database as a feasible solution,
  - g. The optimal solution is found where summation of all the four angle of the cone of friction is minimum,
  - h. The process will continue till it reaches to the end coordinate from starting coordinate.
- ix. End of securing for the stable grasp condition.

### 8. 3D Shape Recognition

The acquisition system consists of a stepper motor on which the object is mounted, one CCD camera and a frame grabber card. The frame grabber card is connected to the PC. The camera is fixed while the object mounted on the stepper motor can rotate around vertical axis with one degree of freedom. The minimal rotation angle of the stepper motor is  $13.85^\circ$  and in this way a total number of 26 images were taken. The photographic view of the image acquisition system is shown below in Fig.26.



Fig. 26. Photographic view of the imageacquisition system

These images are then calibrated and processed through image processing software (Sherlock-32 (Sherlock)). The peripheral points are found out from the calibrated images after thresholding and a database has been made to store the peripheral points of each of the twenty six calibrated images. By CATIA these peripheral points have been employed to regenerate the virtual object, which can be rotated with six degrees of freedom during pattern matching. The first image with the corresponding database points are shown in Fig.27.

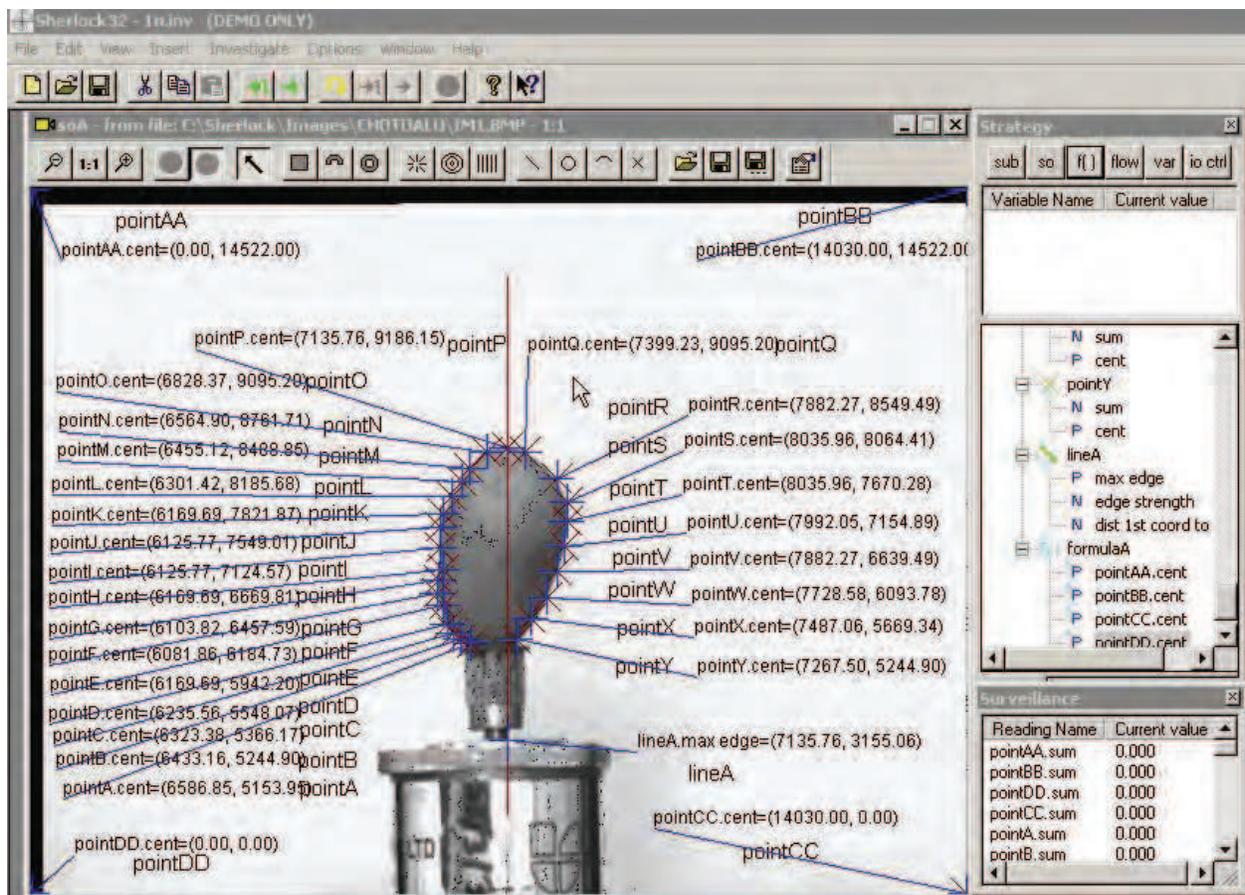


Fig. 27. The Calibrated Image for the object at 1<sup>st</sup> position

The Wire Frame Model of the Object in CATIA is shown in Fig.28. The Online 3-D modeling in CATIA has been shown in Fig.29 and the virtual 3-D synthetic models of the object after rendering in CATIA as viewed from different perspectives are shown in Fig.30 below.

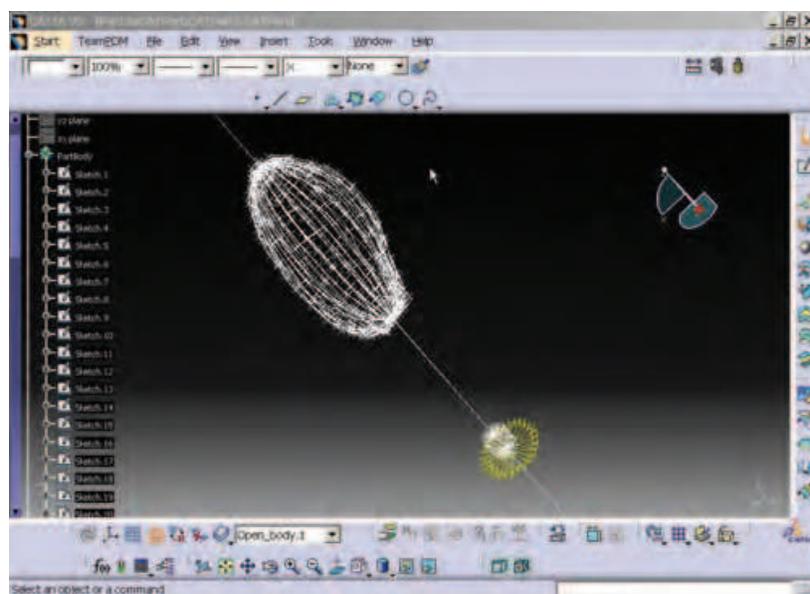


Fig. 28. Wire frame model of the object in CATIA

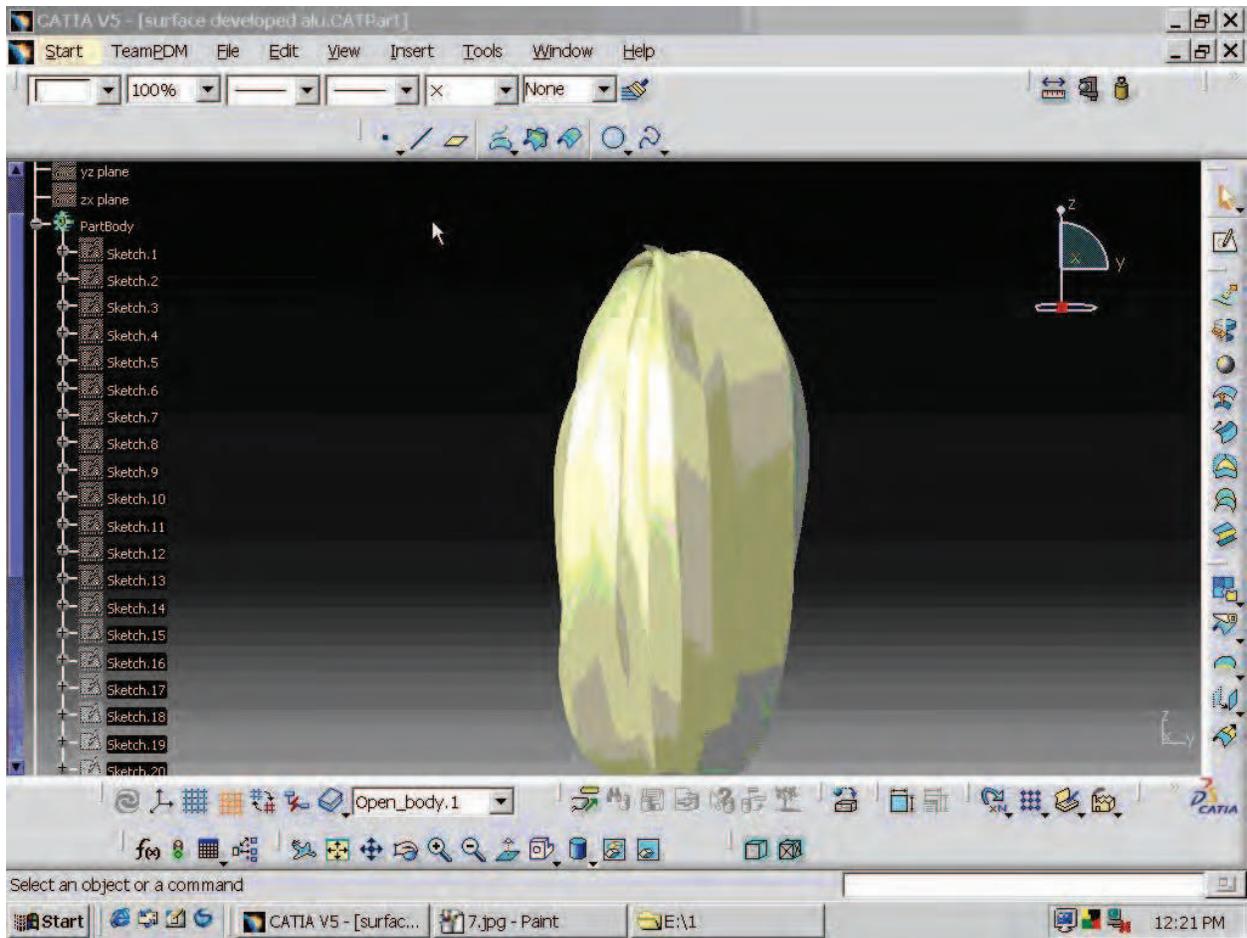


Fig. 29. Online 3-D modeling in CATIA



Fig. 30. 3-D synthetic models of the object after rendering in CATIA as viewed from different perspectives

## 9. Conclusion

From the study of the precision of the robotic hand, the following general conclusions may be drawn:

- a) Object shapes and dimensions have been analyzed through PC-based 2D image processing technique and the information about the object shape has been incorporated in the three-tier hierarchy of the control system for determining the preshape of the fingers during manipulation.
- b) A CCD camera has been employed to grab the image of the object to be gripped and a vision-based algorithm has been developed for analyzing the stable grasp positions and orientations.
- c) It has been found that the vision assistance can minimize the down time for grasping an object in a pre-defined manner by preshaping the fingers.

## 10. Future Scope

- a) In the present setup the hand was so developed that the fingers only ensured planar grasp, though for real time grasping the spatial considerations are to be taken of.
- b) A single CCD camera has been employed for generation of images which resulted in acquisition of dimension of the object in a plane. In order to have the the 3D shape information it would have been better to use another camera at 90°.
- c) 3-D shape information has given a better idea about the object and its grasping strategies. The algorithm, which has been developed for 2-D, may be extended for 3-D grasp.

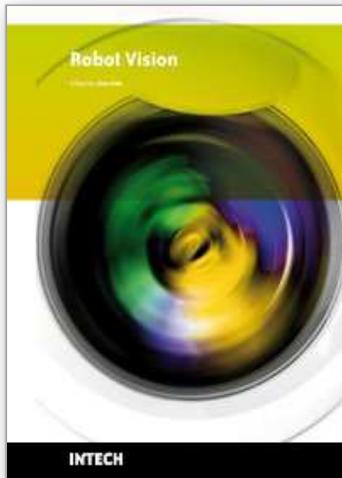
## 11. References

- Geisler, W. (1982). A Vision System for Shape and Position Recognition of Industrial Parts, *Proceedings of International Conference on Robot Vision and Sensory Controls*, Stuttgart, November 1982, Germany
- Marr, D. (1982). *Vision-A Computational Investigation into the Human Representation and Processing of Visual Information*, W. H. Freeman, San Francisco, 1982
- Faugeras, O. D. (1993). *Three-Dimensional Computer Vision: A Geometric Viewpoint*, MIT Press, ISBN 0-262-06158-9, Cambridge, MA
- Horand, R.; Mohr, R.; Dornaika, F. & Boufama, B. (1995). The Advantage of Mounting a Camera in a Robot Arm, *Proceedings of the Europe-China Workshop on Modelling and Invariant for Computer Vision*, pp: 206-213, Xian, 1995, China
- Hartley, R. I. (1994). Self-Calibration from Multiple Views with a Rotating Camera, *3rd European Conference on Computer Vision*, pp. 471-478, Springer-Verlag, Stockholm, Sweden

- Pajdha, T. & Hlavac, V. (1999). Camera Calibration and Euclidean Reconstruction from Known Translations, *Conference on Computer Vision and Applied Geometry, Nordfjordeid, 1<sup>st</sup>-7<sup>th</sup> August 1995, Norway*
- Sonka, M.; Hlavac, V. & Boyle, R. (1998). *Image Processing, Analysis, and Machine Vision*, PWS Publishing, ISBN 053495393X
- Nitzan, D. (1988). Three-Dimensional Vision Structure for Robot Applications, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 10, No. 3, 1988, pp. 291-309
- Victor, A. & Gunasekaran, S. (1993). Range Determination of Objects Using Stereo Vision, *Proceedings of INCARF*, pp. 381-387, December, 1993, New Delhi, India
- Lee, C. S. G.; Fu, K. S. & Gonzalez, R. C. (1994). *Robotics - Control, Sensing, Vision and Intelligence*, Mc-Graw Hill Publication Company, ISBN 0070226253, New York, USA
- Klette, R.; Koschan, A. & Schluns, K. (1996). *Computer Vision-Raumliche Information Aus Digitalen Bildern*, Friedr, Vieweg & Sohn, Braunschweig 1996
- Nishihara, H. K. (1984). Practical Real-Time Imaging Stereo Matcher, *Optical Engineering*, Vol. 23, No 5, 1984, pp. 536-545
- Pollard, S. B.; J. Mayhew, E. W. & Frisby, J. P. (1981). PMF: A Stereo Correspondence Algorithm Using a Disparity Gradient Limit, *Perception*, Vol. 14, pp. 449-470, 1981
- Tarabanis, K. A. & Tsai, R. Y. (1991). Computing Viewpoints that Satisfy Optical Constraints, *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp: 152-158, Maui, , June, 1991, Hawaii
- Maver, J. & Bajcsy, R. (1993). Occlusions as a Guide for Planning the Next View, *IEEE Transactions on Pattern Analysis and Machines Intelligence*, Vol. 15, No. 5, May 1993, pp. 417-432
- Connolly, C. (1985). The Determination of Next Best Views, *Proceedings of IEEE International Conference on Robotics and Automation*, pp. 432-435, Durham, March, 1985, England
- Szeliski, R. (1993). Rapid Octree Construction from Image Sequences, *CVGIP: Image Understanding*, Vol. 58, No. 1, July 1993, pp. 23-32
- Niem, W. (1994). Robust and Fast Modeling of 3D Natural Objects from Multiple Views, *Image and Video Processing II, Proceedings of SPIE*, pp. 388-397, San Jose, February, 1994, CA
- Whaite, P. & Ferrie, F. P. (1994). Autonomous Exploration: Driven by Uncertainty, *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 339-346, Seattle, June, 1994, Washington
- Pito, R. (1999). A Solution to the Next Best View Problem for Automated Surface Acquisition, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 21, No. 10, October 1999, pp. 1016-1030
- Liska, C. (1999). Das Adaptive Lichtschnittverfahren cur Oberflächenkonstruktion mittels Laserlicht, *Master's thesis, Vienna University of Technology, Institute of Computer Aided Automation, Pattern Recognition and Image Processing Group, Vienna, Austria, 1999.*
- Sablating, R.; Tosovic, S. & Kampel, M.(2003). Next View Planning for Shape from Silhouette, *Computer Vision - CVWW'03*, pp. 77-82, Valtice, February, 2003, Czech Republic
- Lacquaniti, F. & Caminiti, R. (1998). Visuo-motor Transformations for Arm Reaching, *European Journal Of Neuroscience*, Vol. 10, pp. 195-203, 1998.

- Desai, J. P. (1998). Motion Planning and Control of Cooperative Robotic Systems, *Ph. D. Dissertation in Mechanical Engineering and Applied Mechanics, University of Pennsylvania*, 1998.
- Metta, G. & Fitzpatrick, P. (2002). Early Integration of Vision and Manipulation. *Giorgio Metta, LIRA-Lab, DIST - University of Genova, Viale Causa, 13 - I-16145, Genova, Italy*, 25<sup>th</sup> October, 2002.
- Barnesand, N. & Liu, Z. Q. (2004). Embodied Categorization for Vision-Guided Mobile Robots, *Pattern Recognition*, Vol. 37, No. 2, February 2004, pp. 299-312
- Kragic, D. & Christensen, H. I. (2003). A Framework for Visual Servoing. In: *Computer Vision Systems*, J.L. Crowley et al. (Eds.), pp. 345-354, Springer-Verlag Berlin Heidelberg, ISBN 978-3-540-00921-4, LNCS 2626
- Bhaumik, S., Barai, N. C. and Ray, R. (2003). Vision Control Mobile Manipulator, *Manufacturing Technology Today, A Journal of Central Manufacturing Technology Institute, Bangalore*, Vol. 3, No. 5, pp. 12-16
- Nakabo, Y.; Ishii, I. & Ishikawa, M. (2002). 3D Tracking Using Two High-Speed Vision Systems, *Proceedings of International Conference on Intelligent Robots and Systems*, pp.360-365, Lausanne, October 2002, Switzerland
- MSC. VisualNastran 4D 2001 R2, MSC Software Corporation, 2 Mc Arthur Place, Santa Ana, California 92707 USA
- Bepari. B. (2006); Computer Aided Design and Construction of an Anthropomorphic Multiple Degrees of Freedom Robot Hand. *Ph. D. Thesis*, Jadavpur University, Kolkata, India
- S. Datta and R. Ray (2007), AMR Vision System for Perception and Job Identification in a Manufacturing Environment. In: *Vision Systems: Applications*, Goro Obinata and Ashish Dutta (Ed), ISBN 978-3-902613-01-1, pp. 519-540, I-Tech, Vienna, Austria, June, 2007
- Sherlock TM, CORECO Imaging, 6969 Trans Canada Highway, Suite 142, St. Laurent, Quebec, <http://www.imaging.com>

IntechOpen



## **Robot Vision**

Edited by Ales Ude

ISBN 978-953-307-077-3

Hard cover, 614 pages

**Publisher** InTech

**Published online** 01, March, 2010

**Published in print edition** March, 2010

The purpose of robot vision is to enable robots to perceive the external world in order to perform a large range of tasks such as navigation, visual servoing for object tracking and manipulation, object recognition and categorization, surveillance, and higher-level decision-making. Among different perceptual modalities, vision is arguably the most important one. It is therefore an essential building block of a cognitive robot. This book presents a snapshot of the wide variety of work in robot vision that is currently going on in different parts of the world.

### **How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Bikash Bepari, Ranjit Ray and Subhasis Bhaumik (2010). Non Contact 2D and 3D Shape Recognition by Vision System for Robotic Prehension, *Robot Vision*, Ales Ude (Ed.), ISBN: 978-953-307-077-3, InTech, Available from: <http://www.intechopen.com/books/robot-vision/non-contact-2d-and-3d-shape-recognition-by-vision-system-for-robotic-prehension>

**INTECH**  
open science | open minds

### **InTech Europe**

University Campus STeP Ri  
Slavka Krautzeka 83/A  
51000 Rijeka, Croatia  
Phone: +385 (51) 770 447  
Fax: +385 (51) 686 166  
[www.intechopen.com](http://www.intechopen.com)

### **InTech China**

Unit 405, Office Block, Hotel Equatorial Shanghai  
No.65, Yan An Road (West), Shanghai, 200040, China  
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元  
Phone: +86-21-62489820  
Fax: +86-21-62489821

© 2010 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](#), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen