Identification and Control of NAO Humanoid Robot To Grasp an Object Using Monocular Vision

Ankita Ranjan and Utkarsh Kumar
Department of Electrical & Electronics Engineering
BIT, Mesra
Ranchi 835215, INDIA
{ankita1901&himutkarsh}@gmail.com

Vijaya Laxmi and Arun Dayal Udai
BIT, Mesra
Ranchi 835215, INDIA

***Abstract* - NAO is a humanoid robot which is developed by Aldebaran robotics, a French robotic company having its headquarters in Paris. Various works related to monocular vision for hand eye coordination in NAO robot has been done which are explained in this paper. Camera calibration is done to determine parameters defining relationship between reference 3D coordinate system and camera coordinate frame, which are combined with transformation and parameters related to camera. Control of robotic arm is a bit complex, it requires complete dynamics study and its implementation. Though it is a wide area, some of the works are done on its arm control using different methods. Accurate control requires identification of D-H parameters and forward and inverse kinematics of NAO arm which needs to have knowledge about degrees of freedom (DOF). Adaptive Neuro Fuzzy Inference System (ANFIS) is used for inverse kinematics calculation. Future works in this direction could be implementing controllers which would help to reduce the stability problem.**

***Index Terms* - NAO robot, hand eye coordination, camera calibration, D-H Parameters, ANFIS**

I. Introduction

Making robot grasp objects is really a challenging work as it requires control of arm and detecting object precisely and accurately so that the hand reaches to the object in time and back to its position. For successful grasping it should be able to adapt to the environment and should not get deviated by any noises in the surroundings. For this the object size should be well known to the robot so that it can detect that object and also its parameters which would make the grab happen accurately. Monocular vision is a vision in which only one eye is used at a time, NAO robot version 2.1 has top and bottom cameras which limits the vision to monocular vision, since it has top and bottom cameras which are placed at an angle. Monocular vision is used to determine the world coordinates using camera calibration. NAO uses top camera by default at a time. Often used practice for selecting frames of reference in robotic practices is the Denavit-Hartenberg parameters. To write the kinematics equations of a manipulator, it gives a standard procedure. Adaptive Neuro Fuzzy Inference System (ANFIS) is a kind of artificial neural network that combines both neural network and fuzzy logic principles. It uses Takagi-Sugeno fuzzy inference system. In this paper, we have used ANFIS for inverse kinematics calculation.

II. Literature Review

In [1], Peipei proposed an intelligent vision localization system. This system consists of a camera mounted on the top of the head of NAO, which provides world vision and the camera helps to determine the local vision. In this two image spaces are used to accomplish grasping. First one is image space determined by ceiling camera, second one is image space determined by its built in camera. In [2], Puheim proposed a feed forward network controller of robotic arm in which stereo vision cameras placed on the head of NAO uses tracking mechanism. 3D spatial coordinates are found out by stereo vision system and used as input to the feed forward controller. Shaw proposed a system in which a camera is mounted on the robot manipulator to allow its gripper to grasp object accurately with the help of controller which is to direct arm for its movement. The controller is used for controlling the arm movement and a microcontroller in the gripper makes the grasping possible[3]. In [4], Shibata proposed the learning technique applied to robotic arm for hand-eye coordination. Allen proposed tracking and grasping of moving objects by the industrial manipulator. In this 3D position of moving object is calculated from stereoscopic system which uses stereo cameras, it is then smoothed and a non linear filter is used to recover the trajectory parameters which is sent to arm control system. The arm controller updates the joint level servos of the arm. After making the tracking stable, system sends the command to arm to grasp the moving object[5]. In [6], Anglani proposed a paper in which controller is designed in combination with monocular vision system. The work by Aashna Sharma, Ishant and Rahul has highlighted method which first creates dataset of known objects and then object recognition is done in 2D using SIFT. Further the 3D localization is done using monocular camera of NAO. In [8], Muller proposed a system which consists of an object disclosure and grasping mobility procedure. Objects are identified by stereo object identifier and hand mobility tracks

978-1-5090-3239-6/17/$31.00©2017IEEE

gimmicked by A based algorithm. In [9], Schmerling proposed a goal directed learning for hand eye coordination in humanoid robot which uses the visuo motor coordination for learning the body kinematics. Lui proposed a method in which robot kinematical parameters ,the calibration point coordinates and the camera extrinsic parameters relative to the target are used for hand eye calibration equation deduction [10]. In [11], Basu proposed a technique for camera calibration in which no predefined patterns are needed. Two algorithms are proposed in this, of which first one uses three different positions of camera for error calculation and focal length estimation. In the second one, first focal lengths are estimated and then error is calculated. In [12], Lenz proposed a method for calibrating robot.In this, robot makes a movement with camera mounted at the gripper and then image features coordinates are extracted and extrinsic calibration is done. Fadi Dornaika and Radu highlighted two methods for transformations from world coordinate system to robot coordinate system and from hand coordinate system to camera coordinate system. First one, solves for rotations and then for translations. Second one,solves simultaneously for rotations and translations[13]. In [14], Zhuang proposed a Complete and Parametrically Continuous(CPC)model for identification of unknown kinematic parameters. In [15], Hager proposed a system that sets robot manipulator using stereo cameras mounted on the manipulator. Visual servoing is done in this paper. The control law has been implemented that executes tracking and stereo control on a single processor. In[16], He proposed a method in which model identification are performed on Devanit-Hartenberg model of humanoid robot. To optimize trajectory of each joint particle, swarm optimization method has been used. ANFIS controller has been implemented for position control of manipulator [20].

III. METHODOLOGY

For the control of NAO robotic arm, the hand eye coordination has to be done. Camera calibration parameters need to be calculated for establishing relation between 3D coordinate system and camera frame. This is essential for finding location of object for grasping task. D-H parameters of NAO left arm is calculated to find the forward kinematics. (ANFIS) is a kind of artificial neural network that combines both neural network and fuzzy logic principles. It uses Takagi-Sugeno fuzzy inference system. In this paper, we have used ANFIS for inverse kinematics calculation.

A. Camera Calibration

It estimates the parameters of a camera. The objective of camera calibration is to determine parameters defining relationship between reference 3D coordinate system and camera coordinate frame which are combined with transformation, and parameters associated with a camera. Camera calibration parameters are the following:

i. Intrinsic parameters: These are the pixel coordinates of image point with respect to corresponding coordinates in camera coordinate system.

ii. Extrinsic parameters: These parameters define orientation of camera coordinate system with respect to the world coordinate system.

Camera calibration parameters calculation:

The calculation of camera calibration parameters was done using camera calibration toolbox in MATLAB .The following steps were done :

a) First the images were loaded for calculating camera calibraton.
b) Secondly,the grid corners were extracted.
c) After that,calibration was done.
d) The images were reprojected to reduce the pixel error.
e) The extrinsic parameters plot was then obtained .
f) After that recomputation of corners were done in order to reduce the pixel error.
g) Calibration was done and error was analysed.

B. Monocular Vision Positioning

NAO cameras are placed at the top and bottom which has wide angular range, making a bit complex for deter-mining exact position. To deal with this, positioning method in monocular view is applied. World frame describes object coordinates in the real world. Camera frame uses the camera center as its origin. A two dimensional point of image is given by $s = [u \quad v \quad w]^T$ and in world coordinate three dimensional point is annotated by $s = [X_w \quad Y_w \quad Z_w]^T$. The equations which are mentioned are based on the pin hole camera model.M_{in} is camera intrinsic matrix. This transfigures camera framework into image pixel framework. The camera extrinsic matrix, $[R \quad t]$ provides transformation from world frame to the camera frame. R is the rotation matrix and t is the translation vector. z_c is the value of z-axis in the camera framework.

.

$$M_{in} = \begin{bmatrix} Kx & 0 & u0 \\ 0 & Ky & v0 \\ 0 & 0 & 1 \end{bmatrix} \tag{1}$$

$$z_c\, s = [M_{in} \quad 0][R \quad t]\, S \tag{2}$$

$$Kx = \frac{f}{dx}, Ky = \frac{f}{dy} \tag{3}$$

M_{in} is given in equation (1). uo,vo is the prinicipal points. *Kx* and *Ky* are the scaling factors in u and v axes. Here, we have used the camera calibration toolbox of MATLAB to evaluate the camera intrinsic matrix. The images used for calibrating is shown in Figure 1. We acquired the transformation from the world coordinate system to the pixel coordinate system by multiplying camera intrinsic parameter matrix by camera extrinsic parameter matrix as shown in (2). M is defined by (4). By setting Z_w a constant value, we can solve the equationbelow. By solving T and P , we obtained the x_w and y_w.

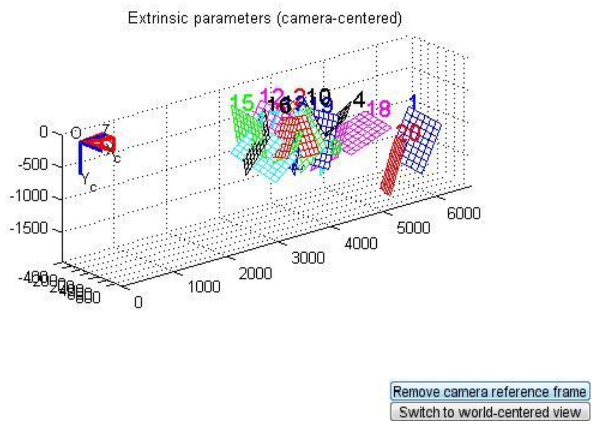
$$M = [M_{in} \quad 0][R \quad t] \tag{4}$$

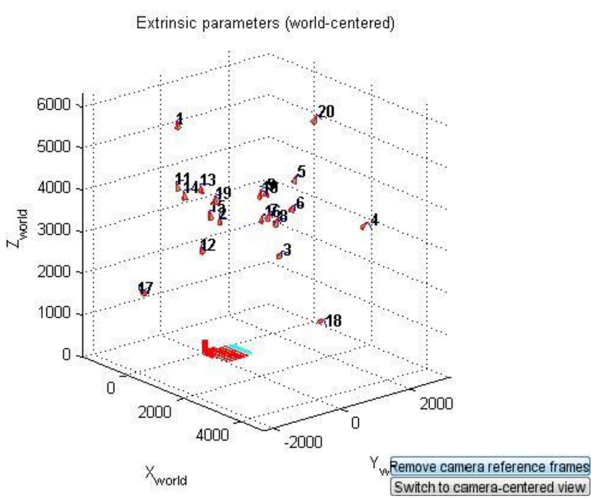
$$M = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \tag{5}$$

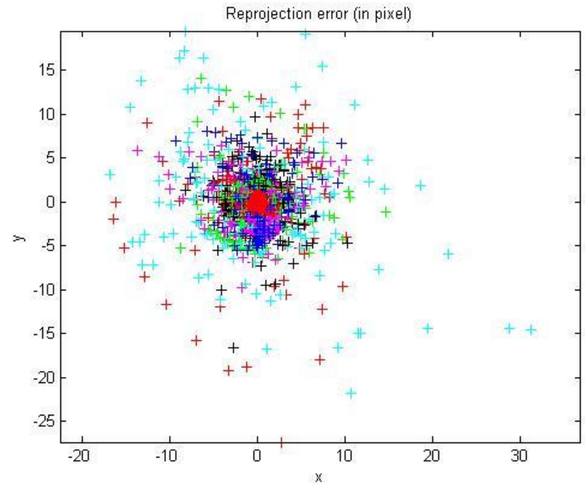
$$T = \begin{bmatrix} m_{11} - um_{31} & m_{12} - um_{32} \\ m_{21} - vm_{31} & m_{22} - vm_{32} \end{bmatrix} \tag{6}$$

$$P = \begin{bmatrix} um_{33} - m_{13} & Z_w + um_{34} - m_{14} \\ vm_{33} - m_{23} & Z_w + vm_{34} - m_{24} \end{bmatrix} \tag{7}$$

$$\begin{bmatrix} x_w \\ y_w \end{bmatrix} = T^{-1} P \tag{8}$$







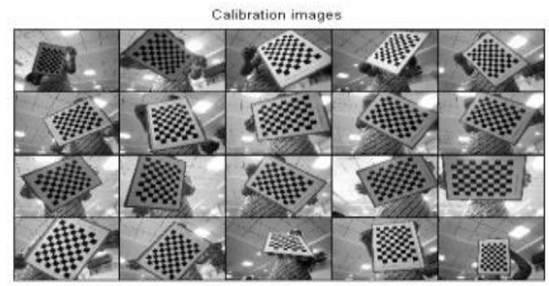


Fig. 1: Calibration results.

C. Robot Kinematics

Study of movement of arms which comprises of multiple degrees of freedom of robot manipulator links is done by applying geometry to its manipulator chains, termed as robot kinematics. Here, in this paper we have NAO robot having 25degrees of freedom in total. Here, we have calculated the D-H parameters of NAO left arm and its corresponding matrices.

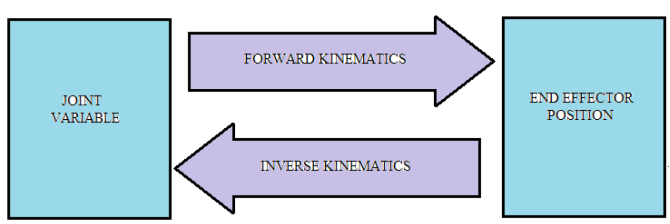


Fig. 2: Depiction of forward and inverse kinematics

1) *Forward Kinematics:* It defines mapping from the joint space to 3D Cartesian space. It finds the end-effectors position and orientation. Forward kinematics is a domain-independent problem. Forward kinematics solution is given by equation (9):

$$T_i = \begin{bmatrix} \cos\theta_i & -\cos\theta_i \sin\alpha_i & \sin\alpha_i \sin\theta_i & a_i \cos\theta_i \\ \sin\theta_i & \cos\alpha_i \cos\theta_i & -\sin\alpha_i \cos\theta_i & a_i \sin\theta_i \\ 0 & \sin\alpha_i & \cos\alpha_i & d_i \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (9)$$

where,

- Joint Offset(d_i) : length of each common normal.
- Joint Angle(θ):rotation about the z-axis.
- Link length (a_i) : the distance on z-axis.
- Twist Angle(α) : angle between two successivez-axes.

D-H parameters calculation:

a. D-H parameters were calculated first analytically.
b. After that, the forward kinematics model was obtained.
c. The results were verified by using Robotics Toolbox.
d. The same results were obtained by writing a MATLAB program for obtaining model.
e. Comparison was made between the model obtained using Robotics Toolbox and without using it.
f. Three Dimensional diagram of the robot arms were obtained.

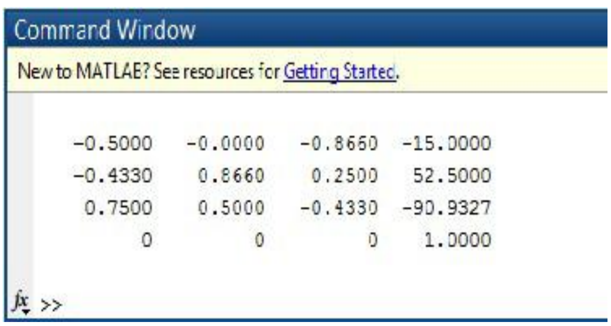


Fig. 3: Calculation of forward kinematics using robotics toolbox.

2) *Inverse Kinematics:* Determines joint variables corresponding to a given end-effector position and orientation. It is much more complicated than forward kinematics problem.

TABLE I: DH Parameters of NAO left arm.

Frame	a(mm)	α(degree)	d(mm)	θ(degree)
LShoulderPitch	0	-pi/2	0	θ_1
LShoulderRoll	0	pi/2	0	θ_2 + pi/2
LElbowYaw	15	pi/2	105	θ_3
LElbowRoll	0	-pi/2	0	θ_4

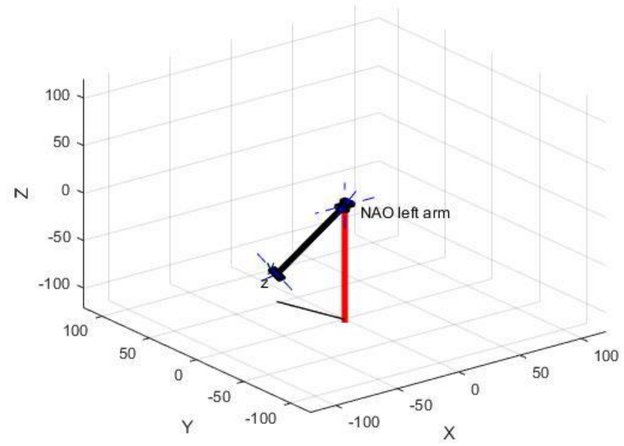


Fig. 4: 3D diagram of the robot.

In inverse kinematics, the transformation matrix is given by equation (9):

$$T = \begin{bmatrix} nx & sx & ax & px \\ ny & sy & ay & py \\ nz & sz & az & pz \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (10)$$

where , p is the translation of end effector from reference frame. n,s,a describes the orientation of end effector and represents x,y,z axes of end effector. After equating transformation matrix to the end effector tool point transformation matrix, the unknown joint angles can be determined. Different joint angles mentioned below:

$$\theta = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \end{bmatrix} \quad (11)$$

The first three columns in the matrices denotes the orientation of the end effectors, whereas the last column denotes the position of the end effectors normal orientation. The inverse kinematics solution of NAO robot left arm are represented from equation (12) to (23) :

$$T = A_{Base}^{0} T_0^1 T_1^2 T_2^3 T_3^4 \, R_z\left(\frac{pi}{2}\right) A_4^{End} \quad (12)$$

$$T' = (A_{Base}^{0})^{-1} T \, (A_4^{End})^{-1} (R_z\left(\frac{pi}{2}\right))^{-1} \quad (13)$$

$$T'' = (T')^{-1} \quad (14)$$

$$\theta_3 = arcsin\left(\frac{T''_{(3,4)}}{l_1}\right) = pi - arcsin\left(\frac{T''_{(3,4)}}{l_1}\right) \quad (15)$$

$$\theta_4 = arccos\left(\frac{l_2 T''_{(1,4)} - l_2 T''_{(1,4)} cos\theta_3}{l_2^2 + l_1^2 \, cos^2\theta_3}\right) \quad (16)$$

$$T''' = T'(T_2^3)^{-1}(T_3^4)^{-1} \quad (17)$$

$$\theta_2 = arctan\left(\frac{T'''_{(2,1)}}{T'''_{(2,2)}}\right) - \frac{pi}{2} \quad (18)$$

$$\theta_1 = arctan\left(\frac{T'''_{(1,3)}}{T'''_{(3,3)}}\right) \quad (19)$$

where,

$$T''_{(3,4)} = l_1 sin\theta_3 \quad (20)$$

$$T''_{(1,4)} = l_1 cos\theta_3 cos\theta_4 + l_2 sin\theta_4 \quad (21)$$

$$T''_{(2,4)} = l_2 cos\theta_4 + l_1 cos\theta_3 sin\theta_4 \quad (22)$$

$$sin\theta_4 = \frac{T''_{(1,4)} + l_2 cos\theta_3 cos\theta_4}{l_2} \quad (23)$$

l_1 = elbow offset
l_2 = upper arm length

IV. ANFIS IMPLEMENTATION

An adaptive neuro-fuzzy inference system is a sort of Artificial Neural Network, based on Takagi-Sugeno fuzzy inference system. Its inference system comprises of a set of IF-THEN rules to be defined for training purpose. Input and output membership functions are defined representing variables, and then rules are applied. Membership functions are of various types such as triangular, Gaussian, sigmoidal, etc. Here, we have used triangular membership function. In this, the neural network has ability to recognise patterns and adapt themselves to adjust with changing environment. Fuzzy-inference system performs decision making. Therefore, it is beneficial to use these techniques in combination with each other. The use of ANFIS helps to overcome shortcomings of fuzzy logic and neural network. Following table shows the input vectors. We have shown few vectors in this table. More than 100 inputs were taken to train ANFIS.

Inverse Kinematics Calculation using ANFIS:

a. First, a number of data sets are used which are given as input to the ANFIS toolbox for training of data,obtained from forward kinematics. Various data sets were used for training for error reduction.
b. ANFIS uses either back propagation or a combination of least squares estimation and back propagation which is known as hybrid for membership parameter estimation.
c. Then, the data was plotted which gave the error in thetraining process. The error was observed and we got the outputdata as the corresponding joint angles.
d. Outputs were obtained in form of surface view by using the surface viewer in the ANFIS toolbox.
e. ANFIS model structure was obtained using number of epochs during training process.

TABLE II: Input parameter vectors

Input vector	1	2	3	4	5
nx	-0.3518	-0.282	-0.5079	-0.7239	0.1619
ny	-0.691	0.3401	-0.2701	0.609	-0.9782
nz	-0.6315	-0.897	-0.818	-0.324	-0.1298
sx	0.2933	-0.847	0.7438	0.6886	-0.9596
sy	0.5593	0.2442	-0.6165	0.666	-0.1254
sz	-0.7754	0.3739	-0.2583	-0.2867	-0.252
ax	0.8889	0.3462	-0.4345	0.0412	0.2302
ay	-0.458	0.9081	-0.7396	-0.4307	0.1653
az	0.0059	0.2354	0.514	-0.9016	-0.959
px	85.0752	42.9728	2.5298	71.3993	-8.5241
py	31.5433	9.7366	-38.364	-77.87	103.9544
pz	54.9294	96.4807	98.8524	-9.4024	-19.2565

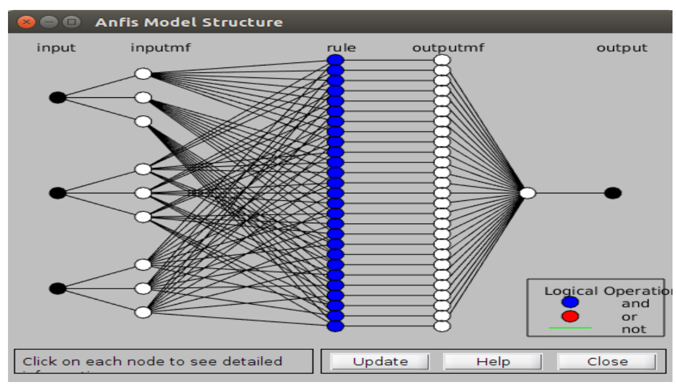


Fig. 5: ANFIS structure.

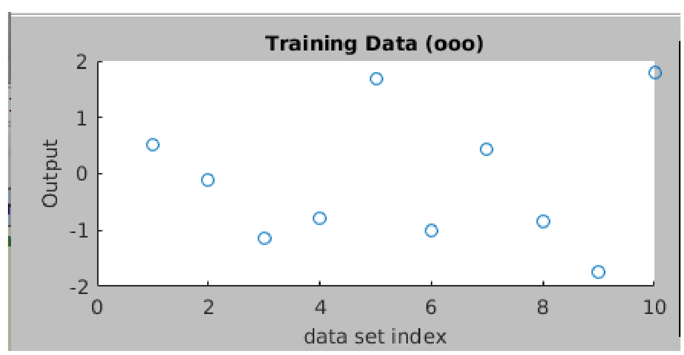


Fig. 6: Outputs obtained from training data.

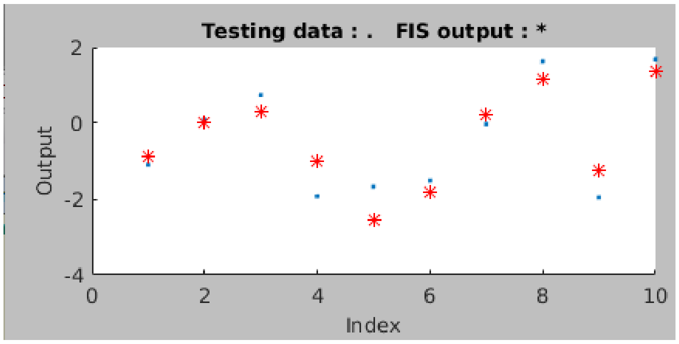


Fig. 7: Outputs obtained from testing data.

Fig. 8: Error obtained while training of data.

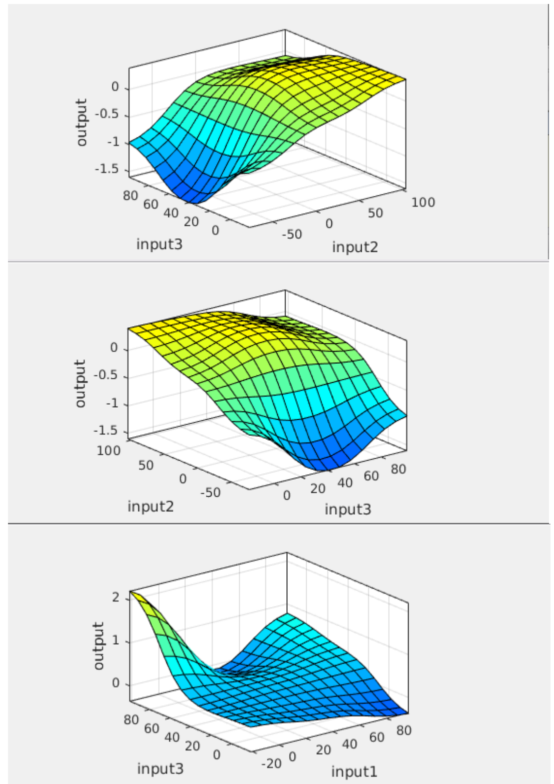
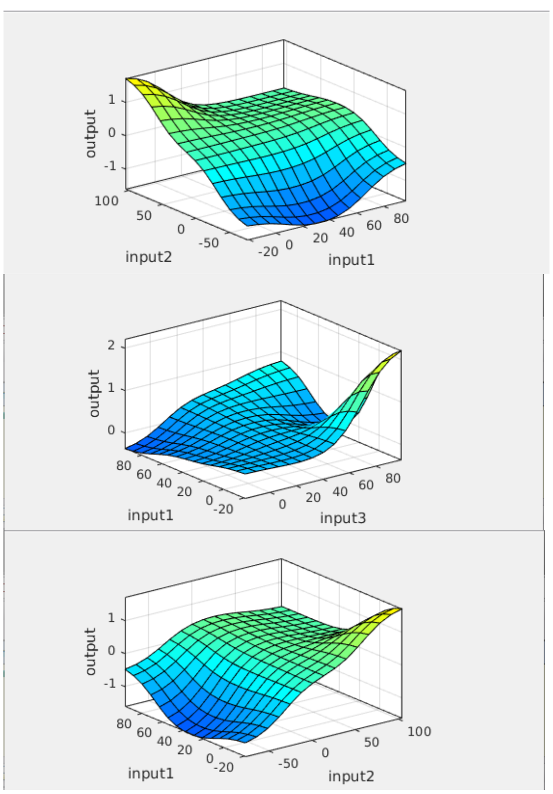


Fig. 9: Surface views obtained in ANFIS.



V. CONCLUSION

This paper mainly focuses on the identification of kinematic model of NAO left arm and calculation of camera intrinsic and extrinsic parameters. Many different approaches are present for camera calibration. Here Bouquets method is used. This paper summarizes various works done on monocular vision in robots and various techniques for control of NAO arms. Adaptive Neuro Fuzzy Inference System (ANFIS) has been implemented for computation of inverse kinematics. For successful training, number of data sets should be large. In this work, the error obtained is2.25×10^{-5}. Future works could include the controller implementation.

REFERENCES

[1] Song Peipei, Li Wenyu,Yan Ningjia, Duan Feng, “Intelligent Vision Localization System of a Service Robot Nao”. *Proceedings of the 34th Chinese Control Conference* , July 28-30, 2015, Hangzhou, China.

[2] Michal Puheim, Marek Bundzel, Ladislav Madarsz, “Forward Control of Robotic Arm Using the Information from Stereo-vision Tracking System”. *14th IEEE International Symposium on Computational Intel-ligence*

and Informatics ,1921 November, 2013 , Budapest, Hungary.

[3] Shaw, K.Y. Cheng, “Object Identification and 3-D Position Calculation Using Eye-in-Hand Single Camera for Robot Gripper*”. IEEE Interna-tional Conference on Industrial Technology (ICIT)*,2016.

[4] Katsunari Shibata, Koji Ito, “Hand-Eye Coordination in Robot Arm Reaching Task by Reinforcement Learning Using a Neural Network”. *IEEE SMC ’99 Conference Proceedings.* Pages:458-463 vol.5, 1999.

[5] Peter K. Allen, Aleksandar Timcenko, Billibon Yoshimi, Paul Michel-
man, “Automated Tracking and Grasping of a Moving Object with a Robotic Hand-Eye System”. *IEEE Transactions On Robotics And Automation*, Vol9, No. 2, April 1993.

[6] Alfred Anglani, Francesco Taurisano, Roberto De Giuseppe, Cosimo Distante, “Learning To Grasp Using Visual Information*”. IEEE Inter-national Symposium on Computational Intelligence in Robotics and Automation*, 8-9 Nov,1999.

[7] Aashna Sharma, Ishant Wadhwa,Rahul Kala, “Monocular Camera Based Object Recognition and 3D Localization for Robotic Grasping*”. Interna-tional Conference on Signal Processing, Computing and Control* ,2015.

[8] Judith Muller, Udo Frese, Thomas Rofer, “Grab a Mug Object Detection and Grasp Motion Planning with the Nao Robot”. *12th IEEE-RAS International Conference on Humanoid Robots,* Nov.29-Dec.1, 2012, Japan.

[9] Mathias Schmerling,Guido Schillaci and Verena V. Hafner, “Goal-Directed Learning of Hand-Eye Coordination in Humanoid Robot*”. 5th International Conference on Development and Learning and on Epigenetic Robotics* ,Aug 13-16, 2015, USA.

[10] Yan Liu, Qinglin Wang, Yuan Li, “A Method for Hand-Eye Calibration of a Robot Vision Measuring System”. *10th Asian Control Confer-ence,*2015.

[11] Anup Basu, “Active Calibration of Cameras: Theory and Implementa-tion*”. IEEE Transactions On Systems, Man, And Cybernetics,* Vol. 2.5, No. 2, February 1995.

[12] Reimar K. Lenz and Roger Y. Tsai, “Calibrating a Cartesian Robot with Eye-on-Hand Configuration Independent of Eye-to-Hand Relationship”. *Proceedings CVPR ’88, Computer Society Conference on Computer Vision and Pattern Recognition,* June1998.

[13] Fadi Dornaika and Radu Horaud, “Simultaneous Robot-World and Hand-Eye Calibration”. *IEEE Transactions On Robotics And Automa-tion*, Vol. 14, No. 4, August 1998.

[14] Hanqi Zhuang, Zvi S. Roth, “A Linear Solution to the Kinematic Parameter Identification of Robot Manipulators”. *IEEE Transactions On Robotics And Automation*, Vol. 9, No. 2, April 1993.

[15] Gregory D. Hager, Wen-Chung Chang, and A.S. Morse, “Robot Hand-Eye Coordination Based on Stereo Vision*”. IEEE Control Systems*, Volume: 15,Issue:1 Pages:3039,1995.

[16] Wei He, Weiliang Ge, Yunchuan Li, Yan-Jun Liu, Chenguang Yang, Changyin Sun, “Model Identification and Control Design for a Hu-manoid Robot”. *IEEE Transactions On Systems, Man, And Cybernetics*, Issue:99,Pages:113,2016.

[17] Dragan Kostic,Bram de Jager, Maarten Steinbuch , Ron Hensen, “Mod-eling and Identification for High-Performance Robot Control: An RRR-Robotic Arm Case Study*”. IEEE Transactions On Control Systems Technology*, Vol. 12, No. 6, November 2004.

[18] S. K. Saha, “Introduction to robotics*”. Tata McGraw-Hill Publishing Company Limited*, 2008.

[19] Fatima Zahra Baghli, Yassine Lakhal, Larbi EI Bakkali, Ouafae Ham-doun, “Design and Simulation of Adaptive Neuro Fuzzy Inference System (ANFIS) Controller for a Robot Manipulator”. *Second World Conference On Complex Systems,IEEE,* Pages:298-303, 10-12 Nov 2014.

[20] Nikolaos Kofinas,“Forward and Inverse Kinematics for the NAO Humanoid Robot”, Chania, July 2012.