Proceedings of the International Conference on Man-Machine Systems (ICoMMS)
11 – 13 October 2009, Batu Ferringhi, Penang, MALAYSIA

Active Stereo Vision Based System for Estimation of Mobile Robot Orientation using Composition Matrix

Paulraj M P[2] Fadzilah Hashim[1] R. Badlishah Ahmad[1] Hema C.R[2] MHA Hamid[2]
[1]School of Computer and Communication Engineering
[2]School of Mechatronic Engineering
University Malaysia Perlis (UniMAP),
Perlis, Malaysia.

***Abstract*- The computation of a mobile robot position and orientation is a common task in the area of computing vision and image processing. For a successful application, it is important that the position and orientation of a mobile robot must be determined properly. In this paper, a simple procedure for determining the orientation of the mobile robot using two cameras is presented. The two cameras are used to capture the images of a mobile robot at various orientations. Four simple neural network models are developed to associate the inputs and output (orientation). First neural network model is used to estimate the orientation of a mobile robot using only the features derived from the center camera. The second neural network model is used for estimating the orientation of a mobile robot using features derived from both the centre and side cameras. The third neural network model is used to estimate the orientation using features derived from the composition matrix. The fourth neural network model is used for estimating the orientation using Singular Value Decomposition (SVD) technique. Simulation results show that the proposed algorithm can be used to estimate the orientation of the mobile robot accurately.**

I. INTRODUCTION

The intelligent mobile robots are widely used in applications such as general indoor and outdoor operations, emergency rescue operations, underwater and space exploration, pipe and duct inspection in power plants, construction environments and so on [1]. Research on mobile robots has attracted much attention in recent years since they are increasingly used in wide range of applications [5-13].

Mobile robots have the capability to move around in their environment and are not fixed to one physical location. For this reason, accurate localization estimation is one of the main requirements for mobile robot navigation. Indoors and outdoors, mobile robots need to know their exact position and orientation in order to perform their tasks [2].

II. BACKGROUND

Different techniques have been used, including several sensors such as sonar sensors, odometry, ultrasonic beacons to obtain a precise position and orientation of a mobile robot [5-16]. Guilherme N. DeSouza et al [3] survey the developments of the last 20 years in the area of vision for mobile robot navigation. Two major components of the paper deal with indoor navigation and outdoor navigation. For each component, they have further subdivided our treatment of the subject on the basis of structured and unstructured environments. For indoor robots in structured environments, they have dealt separately with the cases of geometrical and topological models of space. For unstructured environments, they have discussed the cases of navigation using optical flows, using methods from the appearance-based paradigm, and by recognition of specific objects in the environment. Another survey on the state-of-the-art in sensors, systems, methods, and technologies for mobile robot's positioning is presented by J. Borensteini et al [4].

Accurate position and orientation estimation are extremely important to the successful operation of most autonomous mobile robots. Localization is the process of finding both position and orientation of a vehicle. As a mobile robot moves through its environment, its actual position and orientation always differs from the position and orientation that it is commanded to hold. Therefore, a vision system is always considered as the best sensor to find the current location of the robot in its environment. Antonio Paulino et al [5] present an approach to maintain the positions and orientations of multiple robots using a single camera.

Gijeong Jang et al [6] present a novel localization paradigm for mobile robots based on artificial and natural landmarks. A model-based object recognition method detects natural landmarks and conducts the global and topological localization. In addition, a metric localization method using artificial landmarks is fused to complement the deficiency of topology map and guide to action behavior. The recognition algorithm uses a modified local Zernike moments and a probabilistic voting method for the robust detection of objects in cluttered indoor environments. An artificial landmark is designed to have a three-dimensional multi-colored structure and the projection distortion of the structure encodes the distance and viewing direction of the robot.

Margrit Betke et al [7] describe an efficient method for localizing a mobile robot in an environment with landmarks. They assume that the robot can identify these landmarks and measure robot's bearings relative to each other. Given such noisy input, the algorithm estimates the robot's position and orientation with respect to the map of the environment. The algorithm makes efficient use of their representation of the landmarks by complex numbers. The algorithm runs in time *linear* in the number of landmarks.

James L. Crowley [8] describes a system for dynamically maintaining a description of the limits to free space for a mobile robot using a belt of ultrasonic range devices. These techniques are based on the principle of explicitly representing the uncertainty of the vehicle position as well as the uncertainty inherent in the sensing process. A side effect of matching observations to a local model is a correction to the estimated position of the robot at the time that the observation was made. A Kalman filter update equation is developed to permit the correspondance of a line segment to the model to be applied as a correction to estimated position.

Most of methods, normally, give an estimation of the position and orientation of the vehicle, but, often, they are not able to provide a good estimate of the uncertainty in the measurement. That information is useful in application where multisensor fusion is requested. In contrast to the traditional approach, visual recognition is formulated as one of matching appearance rather than shape. For any given robot vision task, all possible appearance variations define its visual workspace.

S. Noushath, A. Rao and G. Hemanthakumar [9] propose SVD based algorithms for robust face/object recognition and ascertained the efficacy of the SVD based algorithms for both face and object recognition which are useful in robot vision related tasks. Shree K. et al [10] propose the method where a set of images is obtained by coarsely sampling the workspace. The image set is compressed to obtain a low-dimensional subspace, called the eigenspace, in which the visual workspace is represented as a continuous appearance manifold. Given an unknown input image, the recognition system first projects the image to eigenspace. The parameters of the vision task are recognized based on the exact location of the projection on the appearance manifold. An efficient algorithm for finding the closest manifold point is described. The proposed appearance representation has several applications in robot vision.

Goal-oriented navigation of a mobile robot by landmark based techniques is a straightforward and suitable approach. E. Stella, et al [11] method permits to determine the vehicle location and relative uncertainty, when its orientation is obtained by a heading sensor, using a visual landmark based method.

A. L. Betker, et al [13] used a feedforward backpropagation neural network model to estimate the resultant center of mass (COM) trajectory in the sagittal plane. The authors have estimated the COM trajectory for a two-segment inverted pendulum, using clinically available information.

Jason A. Janet et al [14], compare two neural network-based approaches to global selflocalization (GSL) for autonomous mobile robots using a Kohonen neural network model and a region-feature neural network.

The researchers above give us an opportunity to propose our approaches to create a simple feature extraction algorithm and neural network technique to solve the localization problem of mobile robot.

III. Experimental Implementation

A. Image Acquisition

The first camera (center camera C_1) is fixed at the centre of the floor at the height of 2.1m above the floor area. To estimate the orientation of a mobile robot a simple experimental setup is made. The image acquisition system uses two webcameras. The size of the floor area covered by the first camera is 1.7m length and 1.3m width. The second camera (side camera C_2) is fixed at the height of 2.3m above the ground level.

The first camera is fixed at an angle of 90° above the mobile robot and the second camera is fixed at an angle of Θ_2° vertically towards its centroid for increasing field of vision area (Θ_2=22.5°).

179 images of the mobile robot from each webcam at different position and orientation are acquired. Figure1 illustrates the experimental setup used for capturing the images of mobile robot.

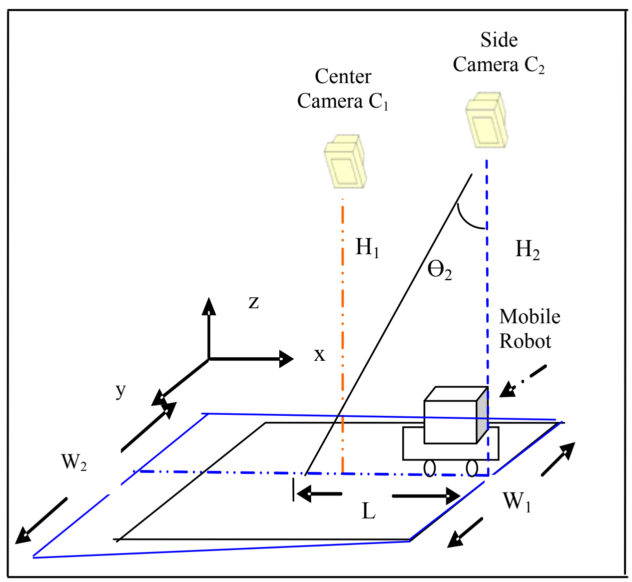
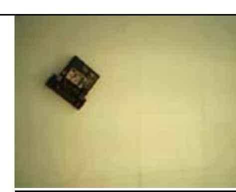


Fig. 1. Experimental Setup

Camera 1	Camera 2
	

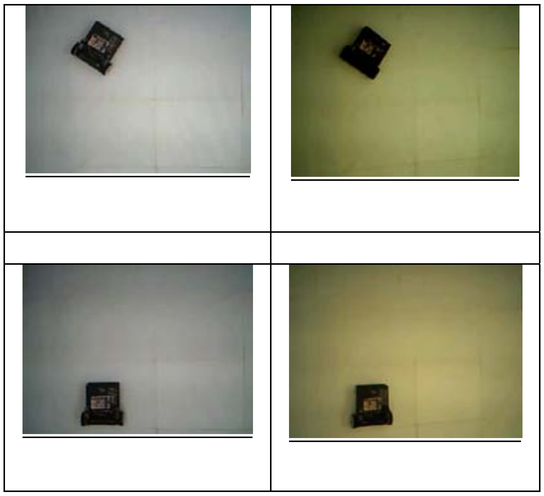


Fig. 2. Samples of images captured at different orientations using two cameras

B. Feature Extraction

The resolution with 640x480 pixel causes considerable delay in the execution of the acquisition process and longer processing time. As a result, all the images are resized to 32x48 pixel before converting into gray-scale images as well as binary images.

Feature Extraction Algorithms for Composition Matrix

1) Resize the original image into 32 x 48 pixel resolution.
2) Convert the resized image into gray-scale image.
3) Convert the gray scale image into binary image by adjusting the threshold value.
4) Representing the resized center camera C_1 images as A and resized side camera C_2 images as B, the composition of both images are obtained by multiplying the image matrix A with the transposed of image matrix B. Hence, the new composition matrix can be represented as AB^T.
5) From the composition matrix, obtain four coordinates to localize the mobile robot, which is rectangular in shape: A= (x_{min}, y_{min}), B= (x_{min}, y_{max}), C= (x_{max}, y_{min}), and D= (x_{max}, y_{max}).
6) Crop the image of mobile robot.
6) From the global image, the global centroids (Gx,Gy), area and perimeter are computed. The local centroids(Lx,Ly), sum of local columns, sum of local rows, moment of local columns and moment of local rows are obtained from the local or cropped image. These parameters are used as input features for training the neural network.

Fig. 4 shows the first three steps involved in feature extraction.

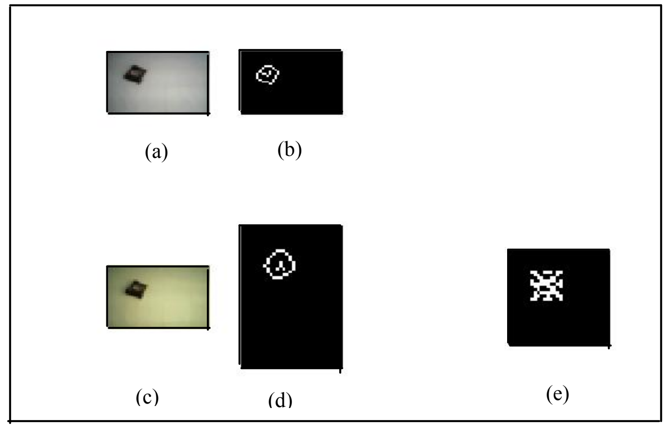


Fig. 4. (a) Resized image from first camera with 32 x 48 pixel, (b) Edge image from first camera, (c) Resized image from second camera with 48 x 32 pixel, (d) Edge image from second camera with transposed matrix (e) Composition matrix from first and second cameras with 32 x 32 pixel.

For each binary image, sum of pixel value along the rows and the columns are all computed. From the computed pixel values, the local region of interest is defined. Fig. 5 shows how the local region of interest is defined from the original image.

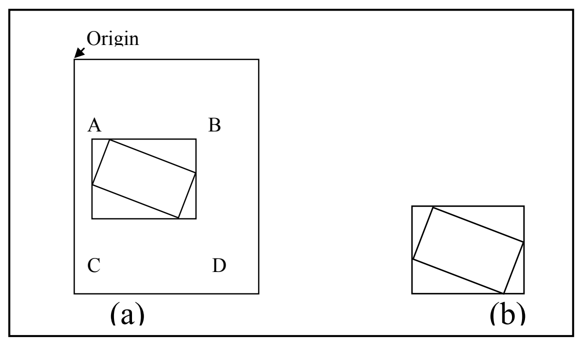


Fig. 5 (a). Global image, (b). Local or Crop image.

For each mobile robot position, the angle of orientation is measured manually.

Feature Extraction Algorithms for SVD Technique

1) Resize the original image into 120 x 150 pixel resolution.
2) Convert the resized image into gray-scale image
3) Convert the gray scale image into binary image by adjusting the threshold value.
4) Representing the resized center camera C_1 images as A and resized side camera C_2 images as B, the composition of both images are obtained by multiplying the image matrix A with the

transposed of image matrix B. Hence, the new composition matrix can be represented as AB^T.

5) Find SVD value of the composition matrx and choose only 14 columns of the singular values. These parameters are used as input features for training the neural network.

IV. NEURAL NETWORK ARCHITECTURE

Artificial Neural Network (ANN) provides alternative form of computing that attempts to mimic the functionality of the brain [20]. In developing the first neural network model, the 11 features obtained from the images of the center camera are taken as the input pattern. The first network model has 11 input neurons, 21 hidden neurons and one output neurons. The 22 features derived from the both camera images are considered as the input features for the second neural network model**.** The second neural network model is designed for estimating the orientation and it has 22 input neurons, 21 hidden neurons and one output neuron. The third neural network model is designed for estimating the orientation using two cameras and composition matrix data features and it has 32 input neurons, 21 hidden neurons and one output neuron. The fouth neural network model is designed for estimating the orientation for SVD technique and it has 14 input neurons, 21 hidden neurons and one output neuron. All hidden neurons have a bias value of 1.0 and the input and hidden neurons are activated by binary sigmoidal activation function of the form

$$f(x) = \frac{1}{(1 + e^{-x})} \qquad (1)$$

Every the neural network models is trained with 130 samples and tested with 179 samples. The initial weights for the above network are randomized between -0.5 and 0.5. A trial weight set consist of 60 sets of randomized weight samples are considered. While training the network models, the performance goal is fixed as 0.001and testing tolerance is fixed as 0.1. Also, while training the network models, the learning rate and momentum factor are chosen as 0.15 and 0.95 respectively.

IV. RESULTS AND DISCUSSIONS

The four neural network models are trained by back propagation procedure with momentum and adaptive learning rate. The training results for the four neural networks are tabulated in Table 1, 2, 3 and 4 respectively.

TABLE I
NEURAL NETWORK TRAINING RESULTS USING CENTER CAMERA

Number of input neurons: 11
Number of Hidden Neurons: 21
Number of output neuron: 1
Activation Function: Binary sigmoidal activation function
Learning Rate: 0.15 Momentum Factor:0.95
Performance Goal: 0.001 qh=1.0 qo=1.0
Testing Tolerance: 0.1
Number of samples used for training: 130
Number of samples used for testing: 179

Trial No.	Mean Classification Rate(%)	Mean Epoch for Training
1	82.26	162235
2	82.26	162235
3	82.26	162235
4	82.26	162235
5	**82.66**	**155837**

TABLE II
NEURAL NETWORK TRAINING RESULTS USING TWO CAMERAS

Number of input neurons: 22
Number of Hidden Neurons: 21
Number of output neuron: 1
Activation Function: Binary sigmoidal activation function
Learning Rate: 0.15 Momentum Factor:0.95
Performance Goal: 0.001 qh=1.0 qo=1.0
Testing Tolerance: 0.1
Number of samples used for training: 130
Number of samples used for testing: 179

Trial No.	Mean Classification Rate(%)	Mean Epoch for Training
1	94.02	**17308**
2	94.11	18252
3	94.28	17340
4	94.18	17514
5	**94.37**	17440

TABLE III
NEURAL NETWORK TRAINING RESULTS USING TWO CAMERAS AND COMPOSITION MATRIX

Number of input neurons: 32
Number of Hidden Neurons: 21
Number of output neuron: 1
Activation Function: Binary sigmoidal activation function
Learning Rate: 0.15 Momentum Factor:0.95
Performance Goal: 0.001 qh=1.0 qo=1.0
Testing Tolerance: 0.1
Number of samples used for training: 130
Number of samples used for testing: 179

Trial No.	Mean Classification Rate(%)	Mean Epoch for Training
1	93.24	10637
2	92.95	**10042**
3	93.07	11292
4	93.29	10467
5	**93.58**	10535

From Table I, it can be observed that the highest mean classification rate is 82.66% and lowest mean epoch is

155837. Next, from Table II and III, it can be observed that the highest mean classification rates are 94.37% and 93.58% respectively. Similarly, it can be observed that the lowest epoch for the second and third neural networks are 17308 and 10042 respectively.

TABLE IV
NEURAL NETWORK TRAINING RESULTS USING COMPOSITION MATRIX WITH SVD TECHNIQUE

Number of input neurons: 11 Number of Hidden Neurons: 21 Number of output neuron: 1 Activation Function: Binary sigmoidal activation function Learning Rate: 0.15 Momentum Factor:0.95 Performance Goal: 0.001 qh=1.0 qo=1.0 Testing Tolerance: 0.1 Number of samples used for training: 130 Number of samples used for testing: 179		
Trial No.	Mean Classification Rate(%)	Mean Epoch for Training
1	77.92	236859
2	78.72	223126
3	78.19	220763
4	78.82	202207
5	78.82	204588

From Table IV, it can be observed that the highest mean classification rate is 78.82% and lowest mean epoch for training the fourth network model is 202207.

Fig. 6 shows the actual values and predicted values of orientations of mobile robot when using center camera.

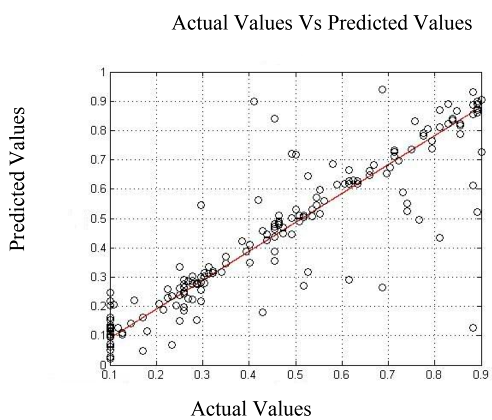


Fig. 6 Actual Vs Predicted Values Using Center Camera

Next, Fig. 7 shows the actual values and predicted values of orientations of mobile robot when using two cameras.

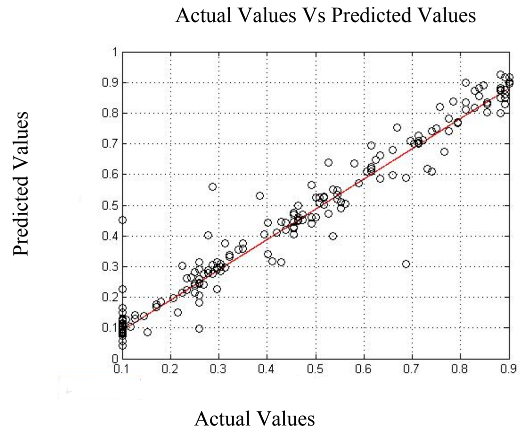


Fig. 7 Actual Vs Predicted Values Using Two Cameras

Fig. 8 shows the actual values and predicted values of orientations of mobile robot when using two cameras and composition matrix.

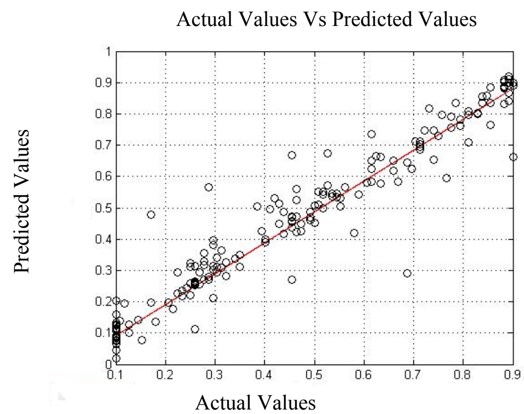


Fig. 8 Actual Vs Predicted Values Using Two Cameras and Composition Matrix

Fig. 9 shows the actual values and predicted values of orientations of mobile robot when using SVD technique.

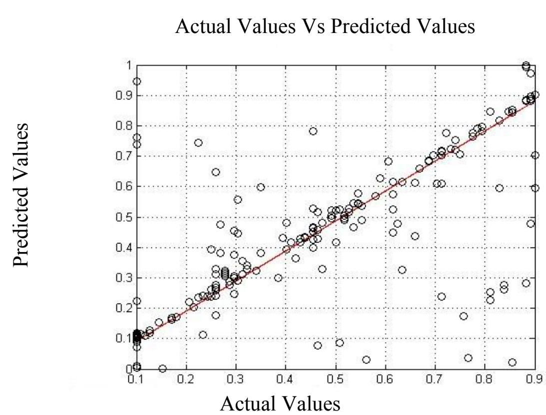


Fig. 9 Actual Vs Predicted Values Using SVD Technique

TABLE V
COMPARISON OF AVERAGE MEAN CLASSIFICATION RATE

Neural Network Model	Average Mean Classification Rate(%)	Average Mean Epoch for Training
Center Camera	82.66	155837
Two Cameras	94.37	17308
Two Cameras and Composition Matrix	93.58	10042
Singular Value Decomposition (SVD)	78.28	202207

From the above results, it is observed that the orientation estimated using the two cameras features provide better results when compared with the single camera results. Even though that the third neural network model (Two Cameras and Composition Matrix) has the lower classification accuracy compared to the two cameras features, but mean epoch for training is the lowest when compared with all the three methods. It is also observed that the orientation estimated using the Singular Value Decomposition (SVD) features provide the results with only 78.82% compared with transposed images, the orientation estimated provide the highest results.

V. CONCLUSION AND FUTURE WORK

A new feature extraction algorithm has been implemented for extracting features from images taken at different orientations. In order to test the proposed features, four simple neural network models are developed for estimating the orientations of a mobile robot. In this paper, two cameras are used to capture images of mobile robot at various orientations between 0° and 90°. In future, it is proposed to include 360° orientations and to apply the proposed method. Further, it is also proposed to improve the performance of the neural network for the accurate estimation of orientation of a mobile robot by increasing the size of the pixel.

REFERENCES

[1] Saeed B. Niku, "Introduction to Robotics: Analysis, Systems, Applications, Prentice Hall, 2001.

[2] Stergios I. Roumeliotis and George A. Bekey, "Distributed Multirobot Localization", *IEEE Transactions On Robotics And Automation,* Vol. 18, No. 5, pp. 781-786, 2002.

[3] Guilherme N. DeSouza and Avinash C. Kak, "Vision for Mobile Robot Navigation: A Survey", IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 24, No. 2, pp. 237-267, 2002.

[4] J. Borenstein1, H.R. Everett, L. Feng, and D. Wehe, "Mobile Robot Positioning & Sensors and Techniques", *Invited paper for the Journal of Robotic Systems, Special Issue on Mobile Robot,* Vol. 14, No. 4, pp. 231-249, 1997.

[5] Antonio Paulino, Jorge Batista and Helder Araujo, "Maintaining the relative positions and orientations of multiple robots using vision", *Pattern Recognition Letters* 22, pp. 1331-1335, 2001.

[6] Gijeong Jang, Sungho Kim, Wangheon Lee, Inson Kweon, "Robust Self-localization of Mobile Robots using Artificial and Natural Landmarks*", Proceedings 2003 IEEE International Symposium on Computational Intelligence in Robotics and Automation,* pp. 412-417, 2003.

[7] Margritt Betke, Leonid Gurvits, "Mobile Robot Localization Using Landmarks", *IEEE Transactions of Robotics and Automation*, Vol. 13, No.2, 1997.

[8] James L. Crowley, " World Modeling and Position Estimation for a Mobile Robot Using Ultrasonic Ranging", *IEEE Conference on Robotics and Automation, ICRA 89,* Vol 3, pp 1574-1579, 1989.

[9] S. Noushath, A. Rao and G. Hemanthakumar, " SVD Based Algorithms For Robust Face And Object Recognition In Robot Vision Application", 24th. International Symposium on Automation & Robotics in Constructions, I.I.T Madras, pp.473-477, 2007.

[10] Shree K. Nayar, Sameer A. Nene and Hiroshi Murase, "Subspace Methods For Robot Vision", *IEEE Transactions on Robotics and Automation,* Vol. 12, No. 5, pp. 750-758, 1996.

[11] E. Stella, G. Cicirelli, A. Distante, "Self-Location of a Mobile Robot with Uncertainty by Cooperation of a Heading Sensor and a CCD TV Camera", *Proceedings of ICPR '96,* pp.303-307, 1996.

[12] Chang-Jia Fang and Shir-Kuan Lin, "A Performance Criterion for the Depth Estimation of a Robot Visual Control System", Proceeding of the IEEE International Conference on Robotics & Automation, pp. 1201-1206, 2001.

[13] A. L. Betker, T. Szturm and Z. Moussavi, "Application of Feedforward Backpropagation Neural Net. To Center of Mass Estimation for Use in a Clinical Environment", - unpublished.

[14] Tien-Fu Lu, Grier C. I. Lin, "An On-Line relative position and orientation error calibration methodology for workcell robot operations," *Robotic & Computer-Integrated Manufacturing*, Vol. 13, No. 2, pp. 89-99, 1997.

[15] Jason A. Janet, Ricardo Gutierrez,Troy A. Chase, Mark W. White and John C. Sutton, Autonomous Mobile Robot Global Self-Localization Using Kohenen and Region-Feature Neural Networks," *Journal of Robotic Sys*., Vol 14, No. 4, pp. 263-282, 1997.

[16] S.N. Sivanandam, M Paulraj "Introduction to Artificial Neural Networks", *Vikhas Publications Company Ltd*. India, 2003.