

# A comparative study on the performance of neural networks in visual guidance and feedback applications

Abhilash T. Vijayan, Ashok Sankar & Sudheer A. P.

To cite this article: Abhilash T. Vijayan, Ashok Sankar & Sudheer A. P. (2017) A comparative study on the performance of neural networks in visual guidance and feedback applications, *Automatika*, 58:3, 336-346, DOI: [10.1080/00051144.2018.1437683](https://doi.org/10.1080/00051144.2018.1437683)

To link to this article: <https://doi.org/10.1080/00051144.2018.1437683>



© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 22 Feb 2018.



Submit your article to this journal [↗](#)



Article views: 119



View Crossmark data [↗](#)



# A comparative study on the performance of neural networks in visual guidance and feedback applications

Abhilash T. Vijayan <sup>a</sup>, Ashok Sankar <sup>a</sup> and Sudheer A. P. <sup>b</sup>

<sup>a</sup>Department of Electrical Engineering, National Institute of Technology, Calicut, India; <sup>b</sup>Department of Mechanical Engineering, National Institute of Technology, Calicut, India

## ABSTRACT

Vision-based systems increase the flexibility of industrial automation applications by providing non-touching sensory information for processing and feedback. Artificial neural networks (ANNs) help such conformities through prediction in overcoming nonlinear computational spaces. They transform multiple possibilities of outcomes or regions of uncertainty posed by the system components towards solution spaces. Trained networks impart a certain level of intelligence to robotic systems. This paper discusses two applications of machine vision. The 3 degrees of freedom (DOF) robotic assembly provides an accurate cutting of soft materials with visual guidance using pixel elimination. The 6-DOF robot combines visual guidance from a supervisory camera and visual feedback from an attached camera. Using a switching approach in the control strategy, pick and place applications are carried out. With the inclusion of ANN to make the strategies intelligent, both the systems performed better with regard to computational time and convergence. The networks make use of the extracted image features from the scene for different applications. Simulation and experimental results validate the proposed schemes and show the effectiveness of ANN in machine vision applications.

## ARTICLE HISTORY

Received 11 July 2017  
Accepted 30 January 2018

## KEYWORDS

Machine vision; industrial automation; visual servoing; ANN; condition number

## 1. Introduction

Industrial applications of robotics and automation rely on the data from its environment for the generation of control signals in decision-making and characterization. Machine vision is a revelation in this domain which provides useful information about surroundings and targets. It caters to a variety of applications varying from simple guidance to continuous feedback. This paper discusses at length, two practical applications involving vision systems. Soft computing techniques impart intelligence to such systems. ANN is an excellent choice with its ability to approximate over the multiplicity of outputs, regions of ambiguity and decision borders.

Machine vision has replaced traditional visual inspection and quality control through creative and contemporary solutions. An overview of strategies, applications, and tools for the industrial vision systems are available in [1]. These are application specific systems, meaning that no single entity handles all the tasks in every field. Each one is designed based on the needs, where the information retrieved is translated into measurements leading to resolutions. The industrial applications are many in this field – starting from part identification, process control, automated visual inspection and robot guidance falling into different levels based on data processing, access and control [2]. Smart sensors with standard or customized solutions replace conventional cameras due to technological

advancement [3]. The design of external lighting for low-level image processing is essential [4] for robust operation even with a fast processor. Vision-based intelligent systems find their place in a variety of applications like packing irregular two-dimensional (2D) shapes [5], vehicle lot management [6] and sealing ports of battery lids [7]. Vision assists versatile operations on different platforms. Online process control in manufacturing [8] uses MATLAB<sup>®</sup>, flexible automatic assembly [9] uses LabVIEW, whereas Open CV and Visual Studio form the platform for the recognition and location of solar panels [10] from aerial images.

Machine vision applications are not only limited to inspection and guidance but also find great extensions in robotics and automation. Feedback control of robotic systems utilizing camera is termed visual servoing which is a flexible technique for closed loop operations. Vision sensor can be either fixed to watch the scene or attached to the robot. The former is called eye-to-hand camera having a global but less precise sight, and the latter is termed eye-in-hand camera with a local but explicit view. The controller design gives two basic approaches to visual servo control [11]: Image-Based Visual Servoing (IBVS) and Position Based-Visual Servoing (PBVS). IBVS utilizes 2D image feature error along with the depth information for deriving the velocity commands with a proven local asymptotic stability. Coupling of translational and rotational components of

camera velocity can lead to unnecessary and complicated camera motions [12]. Far away objects, image Jacobian singularities evoked by specific targets and camera retreat make it difficult especially in physically limited and hazardous environments. PBVS uses 3D pose error and offers globally feasible trajectories. It suffers from servoing failure on many occasions as the tracked object leaves the field of view. It requires a full reconstruction of the 3D environment and pose estimation which is computationally extensive. Inherent shortcomings of these conventional approaches lead to hybrid control laws [13–15] which combine image information and pose vector components. Partitioning the control [16] or switching between controllers [17] also can ensure convergence. The cooperation of eye-in-hand and eye-to-hand cameras is possible for global and local guidance [18,19].

The size of the data obtained from the camera makes it computationally protracted. It causes a mismatch in the speeds of the responses of the robotic and vision system. Due to nonlinear components, the traditional mathematical models involved in decision-making from image data may be ponderous on many occasions. In such cases, ANN provides solutions with approximations nearing accurate models. These networks consist of a connected parallel structure like the human neural system wherein a weighted sum of inputs and bias transforms the output with the help of an activation function. Neurons arranged in layers process the data starting from the input layer to output layer through the hidden layers and adapt their weights. Learning usually happens from known input–output combinations so that ANN can relate the problem data to appropriate solution spaces enabling them to cater to different levels of applications. Multilayered perceptron (MLP) neural networks with sigmoid transfer function is a structure proven for its universal approximation ability. A disjoint decision region developed from a feedforward network with one hidden layer and an arbitrary activation function [20] is capable of approximating any continuous function with appropriate assumptions. The number of hidden layers, its size and type of activation function affect the performance of the network [21].

The generalization ability and fault tolerance of the networks provide intelligent systems utilizing vision in industries in a variety of field applications not only in engineering but also in medical, agricultural and statistics. Object sorting [22], smart transport systems [23], automatic number plate detection [24] and fruit classification [25] use ANN as their think-tank. Computer-aided diagnostics use image analysis with the help of neural networks for an abnormal breast image [26] or in differentiating between pathological and healthy brain images [27].

Challenges due to uncertainty and complex functions in engineering when modelled with soft

computing techniques give fair results. Adaptation capability of neural networks when combined with the continuous feedback of visual servo systems increases their resilience. Possibilities of ANN in visual servoing are numerous- including approximating image Jacobian [28], adjusting to calibration errors and geometrical changes [29], predicting the desired pose [30] and learning control in an unknown environment [31]. Inter-related visual kinematic relations connecting the image and robot movements can be modelled and mapped with ANN [32] and fuzzy-neural networks [33]. Neural networks provide a platform for machine learning in visual servoing with extreme learning machine capability [34] or reinforcement learning [35] where the system should take action for a cumulative reward in a behavioural manner.

As processes and systems get automated, vision being highly accurate and non-touching in sensing, has a special recognition in control applications. Intelligent systems derived through machine learning yield solutions without being explicitly programmed. This paper, motivated by the possibilities of combining machine vision and machine learning, will present two such applications. An automatic soft material cutter with 3 DOF is designed to use the visual information for making an accurate cut through the job. System performance compared after imparting some level of intelligence by incorporating ANN shows considerable improvement. The second application involves both visual guidance and feedback. The 6-DOF robot has to reach the target for pick and place applications. A supervisory camera guides the end effector until the object is in the field of view of the eye-in-hand camera, from where the control law changes. In case of poor system condition, there can be multiple switches for control. ANN is introduced to predict a pose for switching and the modified performance with simulation and experiments indicate the validity of the approach. The results expound the usefulness of ANN in the field of automation, especially when handling large data as in image processing.

The paper addresses the application of neural networks in visual guidance and servoing as follows. Section 2 elaborates the operation of an intelligent robotic cutter and the effect of using ANN. Part 3 explains a supervisory approach in visual servoing implemented on a 6-DOF ABB make robot [36]. The system performance is investigated after including a neural network. Simulation and experimental results demonstrate the effectiveness of the strategy. Section 4 presents the discussions and conclusion.

## 2. Intelligent robotic assembly for soft material cutting

Visually guided systems find their place in many industrial applications. This section will introduce a

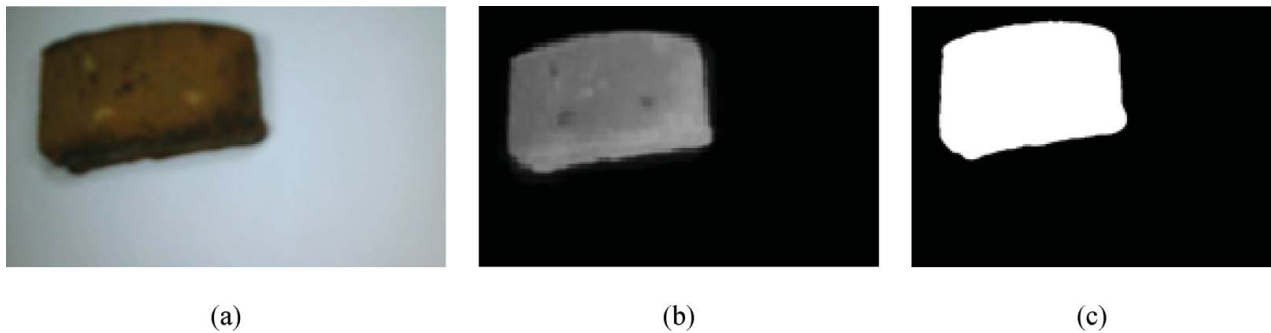


Figure 1. Images of the soft material.

robotic assembly designed to provide accurate and accomplished mechanical cutting of soft materials. The machine will automatically detect and cut the object placed on the platform as per the request. Image-processing operations provide data for computing the signals for robot arm movement and cutting. Artificial neural networks confer a level of general decision-making ability to the system, and a comparison of the performance will follow. Significant components of the system include a vision system, image processor, 3-DOF robotic assembly and the circuits for interface and drive.

### Vision system

A digital camera (Microsoft Lifecam 8 MP) mounted on top of the assembly grabs the frame for further processing. Proper lighting ensures satisfactory results in image processing. Selection of luminary, its distance to the region of interest, light radiation angle, ambient light and glare affect the lighting design. LED light provides an adequate solution for the given system considering the vicinity of the camera to the target and type of application. MATLAB<sup>®</sup> collects the visual data for further processing through a universal serial bus.

### Image processing

Pre-installed driver allows the detection of the camera in the processor as well as MATLAB<sup>®</sup>. A video class input variable containing camera cues stores the frame details. 4D information collected by the camera comprises of R, G, B components of the image and the frame count. A grayscale image containing intensity levels varying between 0 and 255 decreases the size of the data. Background subtraction and transformation provide the Intensity image of the target. Camera quality, lighting conditions and external factors like vibration of the system may degrade the information contained in the frame. De-noising with suitable filters and enhancement algorithms ensures the quality of the image. Many filters are available but a median-filter is very useful when accomplishing noise reduction with edge preservation. Edge detectors provide the shape of

the object. Different morphological operations identify the textural features from the generalized governing equations, portraying the item like a picture with distinct characteristics. Edges define the boundary of the object. Canny method of edge detection is employed as it is more immune to noise. It is likely that the edges are prone to noise as the soft material used is a cake piece. For transforming an image with a grayscale to a binary one, a threshold is specified which determines the boundary for differentiating the black and white pixels which are to follow. Default “graythresh” available in the image-processing toolbox is selected to determine the level. Figure 1 shows the image (a) of a piece of cake used for the experimental evaluation. The fixed camera enables easy background subtraction of the grayscale image for extracting the object (b). Binary version of the image is available (c) for calculating the connected components data like area, perimeter and eccentricity.

### Robotic assembly

A 3-DOF robotic structure capable of providing a precise deep cut in any  $X$ - $Y$  direction in the workspace forms the experimental station. It has a movable arm carrying the cutting blade and a rotating platform holding the soft material. An initial design with Pro-E (Figure 2) allowed a structure for the implementation and the final structure for experimentation is shown in Figure 3. Mounting the camera and light source on top ensures that no shadow or glare falls on the image captured. The tilt free platform is supported by castor wheels and driven by a 12 V, 6 kgcm stepper motor assuming a maximum load of 600 g. Another stepper motor 12 V, 4 kgcm drives the arm carrying the blade and a 5 V and a 1 kgcm servomotor provides the cut. H-bridge-based circuits operate the motors in both directions.

The vision-guided machine cuts the given soft material to requisite weight utilizing standard image-processing algorithms and calculations. The results include a distance “ $r$ ” through which the arm has to move and an angle “ $\theta$ ” by which the platform has to rotate. The nature of the material chosen as the target

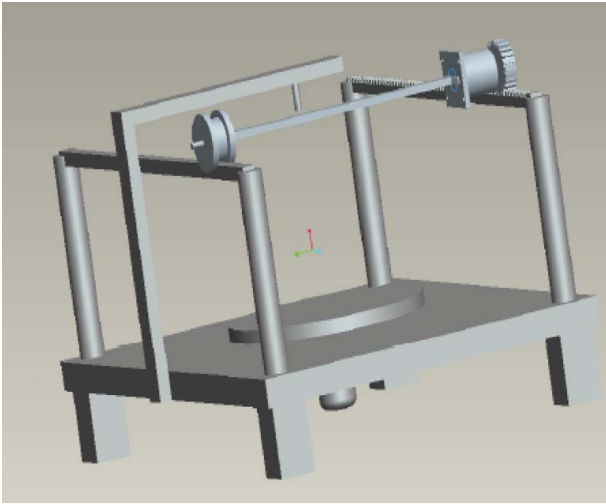


Figure 2. Initial design with Pro-E.

relieves one of the calculations involved in finding the necessary force for the cut.

### ANN for prediction

For a given requisite weight, there can be more than one solution which offers a proper cut. For homogeneous materials, the image area in pixels directly relates to their weight. Hence eliminating the pixels in a straight line can fetch the solution for a given tolerance. This approach results in a considerable error with non-homogeneous materials and irregular shapes. As a result, the system may fall into infinite computational loops or come up with an inaccurate outcome. Unlike conventional digital computers, soft computing techniques like ANN can handle ambiguity driven problems posed by the irregular shapes of the objects, multiplicity of solutions and decision-making. The parallel structure made of simple processing units called neurons can deliver the experimentally acquired knowledge stored in synaptic weights, through learning for arriving at generalized solutions. A multilayer perceptron (MLP) network with a feedforward topology is employed here.

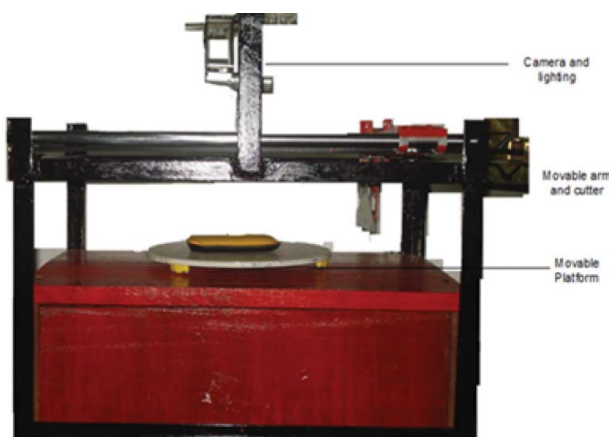


Figure 3. Robotic assembly for soft material cutting.

The design of a neural network which involves determining the hidden layers and its constituent neurons and selection of the activation function is a significant factor affecting its performance apart from the type of system, learning and algorithms. Except for modelling discontinuities and uncertainties, one hidden layer is sufficient for general problems. The inability to model the data will lead to poor fit with the number of neurons in the hidden-layer falling short. Overfitting and large training times are the results of too many neurons, ultimately resulting in inferior outcomes with unseen data. During the study, the number of hidden layer neurons was incremented, starting from 2/3 times the input layer neurons. Different activation functions were employed, and the best results came with “tansig” and two hidden layers while working with non-homogeneous materials. It gave a regression coefficient of 0.9553 compared to 0.9401 with “logsig” during training. Image processing provides distinct features with morphological, textural, dispersion or colour contents. These features define the shape, texture, position and orientation of the workpiece in the workspace. The selected features included descriptors regarding the area, perimeter, edges of the bounding box, centre of mass, eccentricity, major and minor axes of the ellipse encapsulating the image and the pixel intensity values. Requisite weight expressed as a fraction of the actual weight of the object along with 12 selected features forms the input to the network. Two hidden layers have 42 neurons with “tansig” transfer function. Two neurons with a linear transfer function in the output layer predict the distance for the arm movement and the angle for platform rotation, respectively. Experimentally collected results as well as simulations form the training data set of 600 input-output combinations and the training continued till the performance criterion regarding mean square error is reached.

### Performance of the network

For a given error tolerance, direct computational methods falling in infinite loops resulted in a delay to respond. This led to relying on neural networks for probable solutions. Simulation studies proved that results vary with hidden layer size and transfer functions. Table 1 summarizes the results obtained in neural network implementation in comparison with the direct method.

Simulation results show that for regular-shaped and homogeneous materials, direct computation using the image data gives accurate results. Elimination of pixels along straight lines will provide results in some cases; else the procedure repeats for a small increment of angle equivalent to the minimum step of the platform rotation until the tolerance is met. The error will reduce as necessary weight percentage increases. Results are

**Table 1.** Comparison of performance of ANN-based cutting.

S. no.	% Weight	Work piece 1 regular shape; simulation				Work piece 2 irregular shape; simulation				Similar shapes experiment			
		% Error		Time		% Error		Time		% Error		Time	
		D	N	D	N	D	N	D	N	D	N	D	N
1	10	12.1	13.0	8.2	9.6	31.0	12.4	22.3	9.7	25.6	8.2	24.2	25.6
2	20	8.3	7.0	10.3	9.5	10.5	1.5	11.2	9.6	-4.0	6.0	26.6	26.2
3	30	-5.0	1.67	10.8	9.6	-	4.66	-	9.6	8.34	2.5	27.6	26.2
4	40	4.5	2.75	22.0	9.6	-4.5	2.12	24.6	9.6	1.0	4.12	37.2	27.1
5	50	-2.2	3.0	11.4	9.6	9.2	2.4	12.5	9.7	7.6	2.4	28.8	27.0
6	60	1.84	1.8	8.8	9.7	10.34	5.2	16.8	9.8	5.67	1.34	26.6	25.7
7	70	3.57	1.58	11.2	9.6	-	-1.6	-	9.5	-6.8	1.14	29.5	30.6
8	80	-1.2	1.0	12.3	9.6	-3.87	0.85	24.8	9.6	-5.7	-0.6	30.9	26.3
9	90	0.13	1.22	13.5	9.7	2.52	1.34	13.8	9.7	0.45	1.34	29.2	28.6

Note. D, direct method; N, ANN-based method.

similar with ANN for calculation, but image-processing techniques for extracting more features slow down the procedure. For irregular-shaped material, the former method does not provide consistent results. It even failed to exit the computation loop occasionally. Neural networks on the other hand always come up with an outcome, though the error is significant.

Similar shaped soft cake pieces were considered for experimental validation of the strategy. The final cut is affected by the performance of the drive system used for the motors and the nature of the material with regard to homogeneity and dexterity. Hence the error in practical tends to vary from the simulation results. The time shown is indicative of the control which depends on the processor and version of the software. The comprehensive performance of the system gets improved by human-like approach provided by ANN which adds a level of intelligence to the system.

### 3. Supervisory visual feedback control

The previous section addressed the application of machine vision for guidance and control of a Mechatronics system. The drives move over predefined spaces according to the vision system outputs. Feedback is necessary to ensure accuracy in such cases. The use of a camera in the feedback control loop of a robot denotes visual servoing. The following section details a two camera robotic system with visual guidance and feedback for a pick and place application. The camera

attached to the end effector may not see the target inside the workspace for specific configurations (Figure 4(a)). Its local visibility when combined with the global sight of a fixed camera ensures that the target remains in the camera field-of-view. The task is to direct the robot towards the goal with eye-to-hand camera and switch to visual feedback control when the object is fully visible for the eye-in-hand camera (Figure 4(b)).

The experimental station consists of an ABB make industrial robot IRB 1200, a compact short cycle machine with large working envelope and IRC5 (Industrial Robot Controller). It has a payload capacity of 7 kg, an accuracy of 0.02 mm and a reach of 703 mm suiting the needs of material handling and machine tending industries. DH parameters (Table 2) used for deriving the forward and inverse kinematic models of the machine define the robot structure. A pneumatic gripper attached to the robot makes it suitable for pick and place operations. The robot, its workspace, and the target are in the field of view of a fixed camera (Logitech C170, 4 MP) providing visual guidance for the control system. Another camera (Logitech C270, 8 MP), attached to the end effector offers feedback. The target for the experiment is a rectangular block with visible markings so that features can be pre-defined and extracted.

MATLAB<sup>®</sup> programs process the data from the two cameras and calculate the incremental pose required in each step to reach the goal. The robot controller is



**Figure 4.** Robot pose: view from master camera. (a) Initial pose. (b) Intermediate pose.

**Table 2.** D–H parameters of ABB IRB 1200.

Links	$\theta_i(^{\circ})$	$\alpha_i(^{\circ})$	$a_i(\text{mm})$	$d_i(\text{mm})$
1	$\theta_1$	-90	0	342
2	$\theta_2 - 90$	0	350	0
3	$\theta_3$	-90	42	0
4	$\theta_4$	-90	0	351
5	$\theta_5$	-90	0	0
6	$\theta_6 + 180$	0	0	82

programmed in RAPID language and runs parallel with MATLAB<sup>®</sup> during the operation through FTP (file transfer protocol) over Ethernet. FTP server on IRC5 responds to the requests from an FTP client on a remote computer. A graphical user interface, RobotStudio provides the platform for configuration, modelling and offline programming of ABB robots. The path generated through simulation in Robotics and Vision Toolbox in MATLAB<sup>®</sup> is tested in RobotStudio before applying to the robot.

The robot needs to be guided in such a direction that the camera in hand sees the target. Hence, a control law is derived in terms of error in robot end effector pose  $T_e$  and target pose  $T_g$ ;  $T_e = \begin{bmatrix} R(\theta_e) & D_e \\ 0 & 1 \end{bmatrix}$  and  $T_g = \begin{bmatrix} -R(\theta_g) & D_g \\ 0 & 1 \end{bmatrix}$  represent the homogeneous transformation matrix for pose comprising of  $R(\theta)$ , the orthogonal rotation matrix for the orientation and  $D^T$ , the translation vector specifying the coordinates. The control strategy is to minimize the error  $e_T$  given by

$$e_T = T_e - T_g. \quad (1)$$

The components of pose matrix can be stacked to form a vector  $d$  carrying the translational and angular components. A change in the end effector pose  $d$  will follow the error  $e_T$  so that

$$e_d = \dot{d} = J_e V_e, \quad (2)$$

where  $J_e$  is the end effector Jacobian,  $V_e$  is the velocity of the end effector responsible for the motion and  $e_d$  is the error in vector  $d$  between the end effector and the target. The robot will be guided towards the target as  $e_d$  decreases and the end effector approaches its desired pose suitable for the arm to grip the target. A linear controller with a gain  $\beta_1$  will give

$$V_e = -\beta_1 J_e^{-1} e_d. \quad (3)$$

When the features are visible for the eye-in hand camera, it takes over the control. Visible features alone do not guarantee stable operation and hence this is ensured by checking the wellness of the system matrix. Condition Number (CN) represents the ratio of the largest and smallest singular values of the system and is a measure of possible interaction in the system resulting from the translational and rotational components of the velocity. A large condition number

indicates a poorly conditioned system. As the value of CN decreases and approaches unity, the system condition improves. The new goal is to move the end effector towards the target so that the feature error  $e_s$  in the image plane is minimized:

$$e_s = s_{p(x,y)} - s^*. \quad (4)$$

The terms in the error vector carry current values of  $k$  visual features under consideration and their desired values obtained from previously known data or prediction. Camera intrinsic parameters and calibration affect the values of the former whose time variations are related to the camera velocity  $V_c$  in the image plane and the image Jacobian  $J_s$  as

$$\dot{s} = J_s V_c. \quad (5)$$

Similar to the strategy in Equation (2), the control law with a gain factor  $\beta_2$  will be

$$V_c = -\beta_2 J_s^+ e_s. \quad (6)$$

Non square image Jacobian is formed when the number of features selected differs from degrees of robot freedom which leads to Moore–Penrose pseudo-inverse  $J_s^+$  in the computation. The control law in (6) corresponds to conventional image-based visual servoing, where knowledge of the depth is essential for the determination of the Jacobian matrix whose entry for a feature point  $p(x, y)$  and a depth  $z$  is given by

$$J_{s1} = \begin{bmatrix} \frac{1}{z} & 0 & \frac{x}{z} & xy & -(1+x^2) & y \\ 0 & \frac{-1}{z} & \frac{y}{z} & 1+y^2 & -xy & -x \end{bmatrix}. \quad (7)$$

Overall Jacobian is formed by stacking the elements of (7) for the given number of features. This approach is moderately dependable when the switching in control occurs in a region close enough to the target proven for the stability of IBVS. However, the size and shape of this area are yet to be defined which varies with the type of goal, camera parameters, and the robot joint constraints. The wellness of the system matrix indicated by the condition number serves as a qualitative figure of merit for determining the swap of strategies. Loss of features after switching can lead to the transfer of control, back and forth to the master camera or failure of the approach. ANN-based approximations help to overcome the uncertainty associated with the switching region.

### ANN for pose prediction

The supervisory control approach discussed in the previous section has a transfer of control based on the

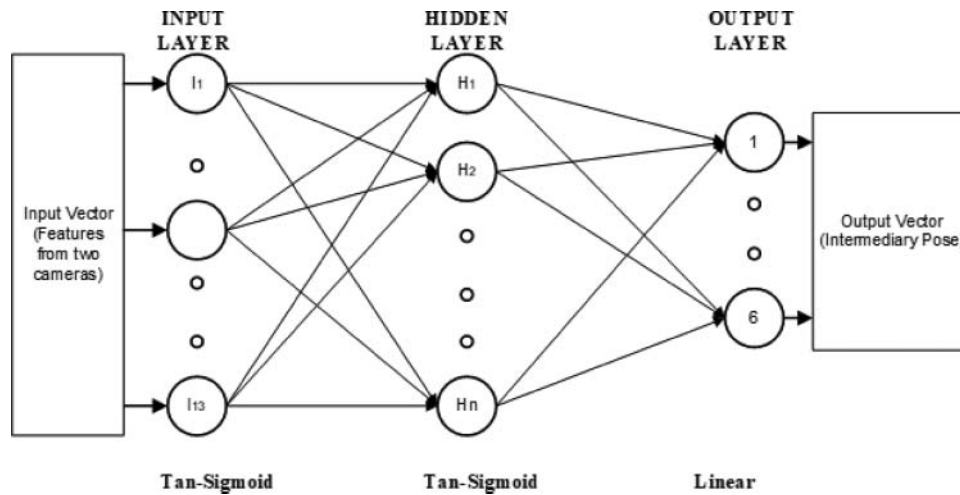


Figure 5. Neural network for predicting the intermediary pose.

wellness of the system matrix. Such a discrete time event can be addressed with the help of a non-conventional control approach like ANN. A trained neural network can predict an intermediary pose for the robot manipulator from any given initial position, favourable for switching to IBVS ensuring the convergence of the method. An MLP network trained by standard back-propagation algorithm forms the think tank of the system. Figure 5 shows the structure of the net with 13 neurons in the input layer. The input vector comprises of visual data collected by the master camera regarding the current pose of the robot and the target along with a distance factor. This scalar represents the point where the camera needs to be placed away from the target for successful servoing. The output layer has six neurons for deriving the predicted pose. ANN is implemented in MATLAB<sup>®</sup> and standard image-processing algorithms are utilized in deriving the image features.

As done in section 2, the number of layers and its size were varied to see the response of the network. The error during training and prediction were noted and the results presented in this section correspond to a neural network with one hidden layer comprising of 52 neurons. The activation function “tansig” performed well in the hidden layer compared with “logsig” during training. The former gave a regression coefficient of  $R = 0.9993$  while the latter had  $R = 0.9990$ . The activation functions in the input and output layers are “tansig” and “linear”, respectively. The database for the training was prepared from both simulation and experimentation results. Out of the 2000 input-output combinations, 15% each was allocated for testing and validation. The supervised learning continued till the set criterion in performance regarding the mean square error norm is reached.

### Performance of the network

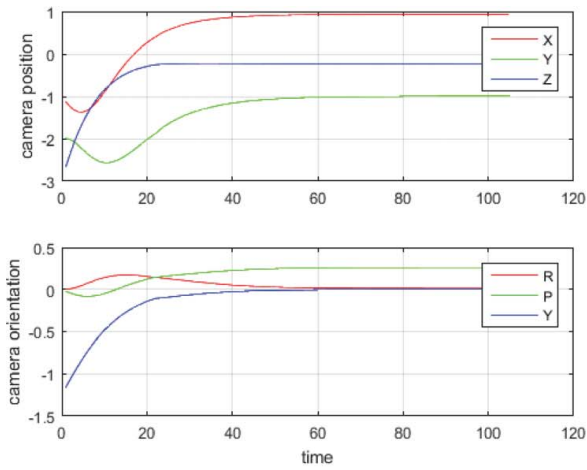
Figure 6 shows the variation of camera position and orientation during the convergence of the two strategies

in simulation (a,b) and experiment (c,d). The targets are different in simulation and real-time evaluation. The control started with (3) under the master camera, and the attached camera took over when the condition of the system matrix improved and the condition number reached the set value of 90. The magnitude of image plane error increases the camera velocity at the point of switching. ANN-based strategy predicts an intermediary pose and the robot reached it with a constant speed and switched to conventional image-based control. For the given case, it converged faster than the former method in simulation. Moreover, the change in velocity at switching is also smooth. Both had similar characteristics during experimentation.

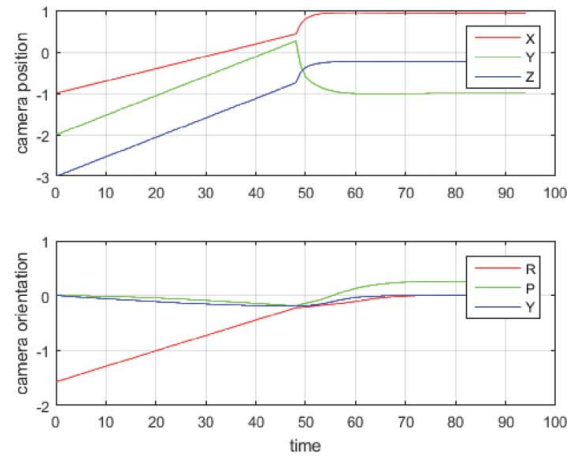
Table 3 summarizes the performance of the network in comparison with the switching scheme based on condition number. The condition number of the Jacobian matrix chosen as the criterion for switching is varied to study its performance. For the given initial conditions, the scheme failed or suffered from multiple switching between control laws for condition numbers above 100 and below 30 selected as the benchmarks for switching. The norm of error in pixels and its distance from the target at the time of transfer of control are noted and tabulated along with the final pixel error at convergence in the table. When the camera approaches the target, the feature points move away from each other in the image plane. Hence the condition of the image Jacobian matrix is improved. The ANN-based strategy gave accurate results and did not fail with permissible limits of the variable input parameter, the proximity to the target. It was varied from 140 cm to 70 cm and the system matrix is fairly conditioned at the time of switching indicating the validity of the scheme. The condition number remained fairly constant in this range of operation.

Switching to control law in Equation (6) alone does not ensure successful visual servoing in all instances. The camera in hand can lose the features and the control switches back to the master camera. Neural

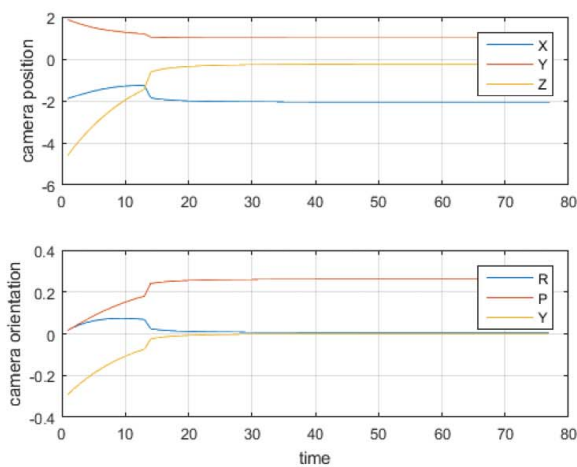




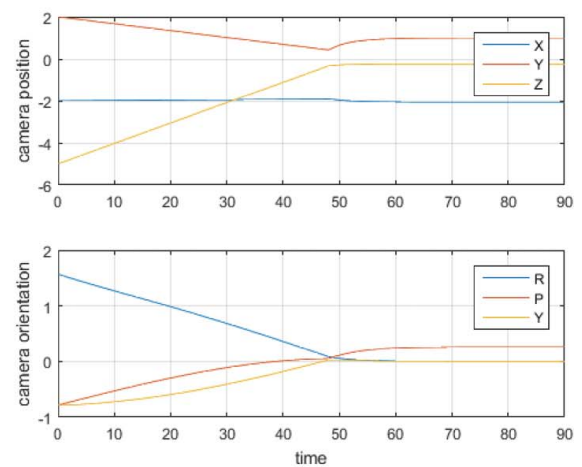
(a) Supervisory control



(b) ANN based Control



(c) Supervisory control



(d) ANN based Control

Figure 6. Camera position and orientation.

network eliminates the event of multiple switching between cameras. The contrast in error between the two control regions causes sudden and wiggly motions, and a variable control gain can suppress this with a compromise on the computation time. However, ANN predicts the intermediary pose suitable for a reasonable change in velocity as shown in Table 4. The maximum change in velocity with ANN is very less compared with supervisory control. It converged faster on all

occasions. The servoing time is dependent on the processor speed and the software version. Since both the strategies go through the same control law at the time of convergence, the overall system performance is dependent on the accuracy of switching. Hence, ANN is superior in this application with regard to time for servoing, failure rate and switching characteristics. It gave smooth switching and prevented the system from falling back to control of master camera.

Table 3. Performance of ANN-based switching scheme I.

Supervisory switching control scheme				ANN-based switching scheme			
Condition number at switching	Proximity from target	Pixel error norm at switching	Norm of final pixel error	Proximity assigned for target	Condition number at switching	Pixel error norm at switching	Norm of final pixel error
100	1.356	2140.4	49.85	1.4	72.14	1900.0	46.42
90	1.152	1999.6	48.49	1.3	62.36	1836.6	46.66
80	1.152	1999.6	48.49	1.2	62.04	1810.9	48.52
70	0.898	1853.6	46.71	1.1	62.80	1858.8	49.02
60	0.898	1853.6	46.71	1.0	61.84	1520.7	48.92
50	0.856	1722.2	49.39	0.9	63.52	1720.4	46.37
40	0.739	1593.6	46.81	0.8	64.77	1978.3	48.83
30	0.670	1365.0	45.80	0.7	68.56	1866.6	47.44

**Table 4.** Performance of ANN-based switching scheme II.

Condition number	Supervisory control scheme			ANN-based control			
	Time (s)	Maximum change in velocity (m/s)	Number of switches	Proximity from the target	Time (s)	Maximum change in velocity (m/s)	Number of switches
100	128	6.5	3	1.4	108	3.1	1
90	108	6.3	1	1.3	102	3.2	1
80	110	4.3	1	1.2	92	3.3	1
70	122	4.1	1	1.1	96	2.1	1
60	126	2.1	1	1.0	102	1.5	1
50	128	0.9	1	0.9	110	0.7	1
40	136	0.8	1	0.8	117	1.8	1
30	142	6.2	3	0.7	108	1.6	1

#### 4. Conclusion

This paper is aimed at studying the usefulness of ANN in machine vision applications. Two separate systems were considered providing specific solutions and the former included a robotic assembly for soft material cutting. The 3-DOF structure automatically detects and cuts the job to the requisite weight. Image-processing techniques allow one to determine the cut by eliminating pixels, given the object is regular shaped and homogeneous. The strategy did not work satisfactorily with irregular shapes and non-homogeneous materials and fell into infinite computational loops. A level of intelligence is required in such problems of ambiguity and decision-making. ANN therefore proved to be pragmatic by providing consistent results. The addition of neural network eliminated the uncertainty and increased the system reliability. The processor speed and the crude mechanical structure affected the final results. Further improvements are possible by increasing the training data and providing feedback. This work considers only feedforward network topology of neural networks. The number of neurons and hidden layer sizes were determined by iterative procedures. In addition, the possibility of replacing the network structure with radial basis function (RBF), recurrent neural networks (RNN) or ANFIS opens the scope for future research.

The second application consisted of a visual servoing system making use of two cameras. For targets initially not visible for the eye in hand camera, the fixed one guides the robot to a pose suitable for switching the control. Loss of features during the operation can lead to failure in servoing. Hence, a check on the system matrix provides the criterion for switching. Numerous switches result from an ill-conditioned matrix. The strategy may fail when the robot blocks the view of the target for the master camera. The multiple input–output nature of the system makes it prone to loop interactions. A higher condition number indicates either a high value of the highest singular value or a low value of the smallest singular value and the latter being a fundamental problem in control causing uncertainty. The addition of a neural network to predict the intermediary pose suitable for stable switching improved the accuracy and performance by


providing a solution for every occasion avoiding multiple switches, converging faster and enhancing system condition. In this case also, the experiment is done with only feedforward network. The presented results gave a consistent output because supervisory control is prone to multiple switching. The study can be further extended by replacing the network structure with RBF or RNN. The choice of proximity assigned is similar to selecting different condition numbers for switching. As the system is liable to interaction resulting from the coupling of translational and rotational components of camera velocity, a lower condition number is preferred at the time of switching. ANN can ensure this in the given experiment. Additional research is possible for a target in motion in the robot workspace.


Problems in engineering are modelled mathematically for probable solutions. However, the size of the data, shape of the object and application require non-conventional approaches for unravelling the uncertainty caused by vision systems. In both the applications presented, specific regions of operation can be identified where accustomed methods prove to be favourable. ANNs can extend the scope of such applications by providing a smarter outlook, especially those involving machine visions. This work highlighted the system reliability attained due to consistent and accurate outcomes provided by ANN. Further possibilities lie in adopting other soft computing techniques like fuzzy and genetic algorithm and their combinations for industrial automation. When a robot operates in parallel with other robots and humans, decision-making regarding the speeds and safety is essential. The multiplicity of similar targets in the field of view demands behavioural approach from robots where reinforcement learning control could be more effective than conventional neural approaches. New learning capabilities can be added to the robot with experience learning which gives the scope for an empirical control zone with the cooperation of two cameras in visual servoing.


#### Disclosure statement

No potential conflict of interest was reported by the authors.

## ORCID

Abhilash T. Vijayan  <http://orcid.org/0000-0001-5818-3696>

Ashok Sankar  <http://orcid.org/0000-0002-1898-0903>

Sudheer A. P.  <http://orcid.org/0000-0003-0644-3702>

## References

- [1] Malamas EN, Petrakis EG, Zervakis M, et al. A survey on industrial vision systems, applications and tools. *Image Vis Comput.* 2003 Feb;21(2):171–188.
- [2] Golnabi H, Asadpour A. Design and application of industrial machine vision systems. *Robot Comput-Integrated Manuf.* 2007 Dec;23:630–637.
- [3] Fernandes AO, Moreira LFE, Mata JM. Machine vision applications and development aspects. 2011 9th IEEE International Conference on Control and Automation; 2011 Dec 19–21; Santiago, Chile; 2011. p. 1274–1278.
- [4] Kopparapu SK. Lighting design for machine vision application. *Image Vis Comput.* 2006 Jul;24:720–726.
- [5] Bouganis A, Shanahan M. A vision-based intelligent system for packing 2-D irregular shapes. *IEEE Trans Autom Sci Eng.* 2007 Jul;4(3):382–394.
- [6] Lin S-F, Chen Y-Y, Liu S-C. A vision-based parking lot management system. 2006 IEEE International Conference on Systems, Man and Cybernetics; 2006 Oct 8–11; Taipei; 2006. p. 2897–2902.
- [7] Gao M, Li X, Yang Y, et al. An automatic sealing system for battery lid based on machine vision. 2017 IEEE 26th International Symposium on Industrial Electronics (ISIE); Edinburgh; 2017. p. 407–411.
- [8] Cheng Y, Jafari MA. Vision-based online process control in manufacturing applications. *IEEE Trans Autom Sci Eng.* 2008 Jan;5(1):140–153.
- [9] Nerakae P, Uangpairroj P, Chamniprasart K. Using machine vision for flexible automatic assembly system. *Procedia Comput Sci.* 2016;96:428–435.
- [10] Yao YY, Hu YT. Recognition and location of solar panels based on machine vision. 2017 2nd Asia-Pacific Conference on Intelligent Robot Systems (ACIRS); Wuhan; 2017. p. 7–12.
- [11] Chaumette F, Hutchinson S. Visual servo control. I. Basic approaches. *IEEE Robot Autom Mag.* 2006 Dec;13(4):82–90.
- [12] Chaumette FCA. Potential problems of stability and convergence in image-based and position-based visual servoing. *The Confluence of Vision and Control Lecture Notes in Control and Information Sciences*; Springer-Verlag; 1998. p. 66–78.
- [13] Chaumette F, Malis E. 2 1/2 D visual servoing: a possible solution to improve image-based and position-based visual servoings. *Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings*; San Francisco (CA); 2000. p. 630–635, Vol. 1.
- [14] Assa A, Janabi-Sharifi F. Hybrid predictive control for constrained visual servoing, 2014 IEEE/ASME International Conference on Advanced Intelligent Mechatronics; Besacon; 2014. p. 931–936.
- [15] Wu CY, Song KT. Hybrid visual servoing design for a continuum robot under visibility constraint and voice commands. 2016 16th International Conference on Control, Automation and Systems (ICCAS); Gyeongju; 2016. p. 1255–1260.
- [16] Corke PI, Hutchinson SA. A new partitioned approach to image-based visual servo control. *IEEE Trans Robot Autom.* 2001 Aug;17(4):507–515.
- [17] Gans NR, Hutchinson SA. Stable visual servoing through hybrid switched-system control. *IEEE Trans Robot.* 2007 Jun;23(3):530–540.
- [18] Flandin G, Chaumette F, Marchand E. Eye-in-hand/eye-to-hand cooperation for visual servoing. *Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings*; 2000 Apr 24–28; San Francisco, CA; 2000. p. 2741–2746.
- [19] Elarbi-Boudihir M, Al-Shalfan KA. Eye-in-hand/eye-to-hand configuration for a WMRA control based on visual servoing. 2013 IEEE 11th International Workshop of Electronics, Control, Measurement, Signals and Their Application to Mechatronics; Toulouse; 2013; p. 1–6.
- [20] Huang Guang-Bin, Chen Yan-Qiu, Babri HA. Classification ability of single hidden layer feedforward neural networks. *IEEE Trans Neural Netw.* 2000 May;11(3):799–801.
- [21] Sheela KG, Deepa SN. Review on methods to fix number of hidden neurons in neural networks. *Math Probl Eng.* 2013;2013:1–11.
- [22] Govender P, Li Z. ANN based intelligent vision and sorting system. 2010 IEEE 17th International Conference on Industrial Engineering and Engineering Management; Xiamen; 2010. p. 394–399.
- [23] Siyal MY. A neural vision based approach for intelligent transportation system. *Industrial Technology, 2002. IEEE ICIT '02. 2002 IEEE International Conference on*; 2002. p. 456–460, Vol. 1.
- [24] Volna E, Kotyrba M. Vision system for licence plate recognition based on neural networks. 13th International Conference on Hybrid Intelligent Systems (HIS 2013); Gammarth; 2013. p. 140–143.
- [25] Zhang Y, Phillips P, Wang S, et al. Fruit classification by biogeography-based optimization and feedforward neural network. *Expert Syst.* 2016 Jun;33(3):239–253.
- [26] Wang S, Rao RV, Chen P, et al. Abnormal breast detection in mammogram images by feed-forward neural network trained by Jaya algorithm. *Fundamen Inform.* 2017 Mar 11; 151.(1–4): 191–211.
- [27] Lu Z, Lu S, Liu G, et al. A pathological brain detection system based on radial basis function neural network. *J Med Imaging Health Inform.* 2016 Sep 1;6(5):1218–1222.
- [28] Stanley K, Wu QMJ, Jerbi A, Gruver W. Neural network-based vision guided robotics. *Proceedings 1999 IEEE International Conference on Robotics and Automation*; Detroit (MI); 1999. p. 281–286, Vol. 1.
- [29] Cupertino F, Giordano V, Mininno E, et al. A neural visual servoing in uncalibrated environments for robotic manipulators. 2004 IEEE International Conference on Systems, Man and Cybernetics; 2004 Oct 10–13; The Hague, Netherlands; 2004. p. 5362–5367.
- [30] Sunami Y, Minami M, Yanou A. Experimental evaluations of prediction servoing to moving object by hand-eye robotic system. *The SICE Annual Conference 2013*; Nagoya (Japan); 2013. p. 1310–1315.
- [31] Nan-Feng X, Nahavandi S. Visual feedback control of a robot in an unknown environment (learning control using neural networks). *Ind Robot: Theory, Model Control.* Jan. 2006. p. 779–798.
- [32] Al-Junaid H. ANN based robotic arm visual servoing nonlinear system. *Procedia Comput Sci.* 2015;62:23–30.

- [33] Gu J, Wang H, Pan Y, et al. Neural network based visual servo control for CNC load/unload manipulator. *Optik – Int J Light Electron Opt.* 2015 Dec;126(23):4489–4492.
- [34] Yüksel T. Intelligent visual servoing with extreme learning machine and fuzzy logic. *Expert Syst Appl.* 2017 Apr;72:344–356.
- [35] Miljković Z, Mitić M, Lazarević M, et al. Neural network reinforcement learning for visual control of robot manipulators. *Expert Syst Appl.* 2013 Apr;40(5):1721–1736.
- [36] ABB IRB 1200 Product Manual. Available from: [https://library.e.abb.com/public/bd8c0f3a077d57e4c1257ceb00722d6e/IRB1200\\_ROB0275EN.pdf](https://library.e.abb.com/public/bd8c0f3a077d57e4c1257ceb00722d6e/IRB1200_ROB0275EN.pdf)