IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received October 10, 2020, accepted November 2, 2020, date of publication November 11, 2020, date of current version November 19, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3037540

A Vision-Based Path Planning and Object Tracking Framework for 6-DOF Robotic Manipulator

ADIL SHAHZAD[1], XUESHAN GAO[2], (Member, IEEE), AWAIS YASIN[3], KAMRAN JAVED[3], AND SYED MUHAMMAD ANWAR[4], (Member, IEEE)

[1]School of Mechatronical Engineering, Beijing Institute of Technology, Beijing 100081, China
[2]State Key Laboratory of Robotics and System, Harbin Institute of Technology, Harbin 150001, China
[3]Department of Computer Engineering, National University of Technology, Islamabad 44000, Pakistan
[4]Department of Software Engineering, University of Engineering and Technology Taxila, Taxila 47050, Pakistan

Corresponding author: Syed Muhammad Anwar (s.anwar@uettaxila.edu.pk)

ABSTRACT Industrial robots are widely used for repetitive, humanly unmanageable, and hazardous tasks. Hence, an improvement in the production efficiency of industrial robot manipulators is of prime concern. This can be achieved through machine vision and path planning techniques with a focus on localization and shortest path calculation. In particular, this is important for manufacturing and bottle filling industries which extensively use robotic manipulators to place/displace bottles during production and post refill placements. This is even more challenging when soft, fragile, or opaque objects have to be detected, since it is significantly difficult for robot vision to focus on their indistinguishable features. To this end, we present an ensemble robot framework with a stereo vision system for tracking colored objects which are sensed using blob analysis. An ensemble robotic framework with neural networks is proposed for predicting and thereby overcoming the inbuilt geometric error present in stereo vision systems. Moreover, we have simplified 2-D correspondence problem to 1-D by using a non-rectified stereo camera model and object tracking by applying the triangulation technique in 3D stereo vision coordinate system (SVCS). Subsequently, the SVCS is transformed into robot stereo vision coordinate system for tracking the object centroid by using an RGB marker placed on the object. Finally, in the learning model we have combined color region tracking with machine learning to achieve high accuracy. The outcomes are in accordance with the designed model and successfully achieve path prediction with up to 91.8% accuracy.

INDEX TERMS Robot stereo vision, path planning, robotic vision, robotic manipulators.

I. INTRODUCTION

Recently industrial robots and robotic manipulators are getting more sophisticated, highly diversified, and designed specifically to achieve a high level of autonomy. This is in comparison to the earlier concepts where robots were considered as mechanical manipulators [1]. Moreover, object detection and tracking for these robotic manipulators has attracted significant attraction with an aim to provide precise control mechanisms [2]. In particular, to make robots flexible, automated object detection has benefited towards a more generalized and adaptive behavior for different objects and positions [3]. This has increased the utility of robotics in industries [4], and enabled complex path designs for various types of objects [5]. Hence, there has been a significant focus on the development of robotic control systems embedded with a vision system. Further, there is a widespread adaptation of machine learning-based methods, where efficient computer vision frameworks are utilized. Moreover, neural networks and deep learning-based methods are also increasing in popularity [6], but care must be taken in terms of explainability and generalization of such models.

For robotic manipulators, the degree of freedom (DOF) is often one of the most important factors for their operation. For instance, for tasks such as welding, pelleting, painting, packaging, grinding, and pick and place jobs, DOF plays a significant role. The DOF ranges from 1-DOF to 6-DOF and is selected depending upon the nature of the work required

The associate editor coordinating the review of this manuscript and approving it for publication was Heng Wang.

to be performed in an industrial environment [7]. In bottling industry, these manipulators have to deal with soft and fragile objects which introduces challenges in efficiency and handling. Herein, we propose an open system environment with a standard 6-DOF industrial STEP SD 700 manipulator to execute the gripping task of soft objects through robot stereo vision (RSV) system. We set up a precise experiment in the laboratory with 6-DOF industrial manipulators equipped with RSV to perform a tracking task for soft objects having red, green, and blue (RGB) color marker in an open environment. Our aim is to enable the displacement task for distinguishable objects. In particular, the color marking technique for objects was employed to simplify and construct correspondence between the stereo camera vision system. Our proposed framework attempts to resolve three main challenges including precise manipulation, real-time calibration, and accurate path planning and object placement. Our major contributions are,

- We design and build a parallel stereo charged coupled device (CCD) camera system, which can accurately track the colored region of the RGB marker. Further, this region of interest (ROI) was separated from image live stream through blob analysis and triangulated for calculation of 3D coordinates in the world coordinate frame. The marker's 3D coordinates were transformed into the robot frame of axis (robot stereo vision coordinate system) and used for evaluation of robot path and trajectory planning. Hence, our proposed model estimates the distance in real-time the robot has to move along the three axes to pick and place an object.
- The RGB marker centroidal data were compared in real-time by calibrating 6-DOF industrial robot movements in the 3D world frame using the RSV system. The RSV system was trained to perform RGB marker tracking from a distance of 10 meters within a millimeter range.
- We incorporated an ensemble technique to improve the performance in predicting output trajectories. To this end, we propose a framework which combines both unrectified RSV outcomes and nonlinear principal component analysis-neural network (NLPCA-NN) in a single module using the combiner circuit (ensembler). This enabled us in achieving accurate path planning for object placement with high accuracy.

II. RELATED WORKS

Robot vision incorporates advanced techniques and algorithms both from computer vision and machine learning domains. This mainly facilitates robot kinematics through reference frame, and vision sensor calibration which affects the physical abilities of a robot and its performance in an environment. Visual servoing is an appropriate example of a technique which can only be termed as robot vision and not computer vision in general. It explicitly involves controlling the motion of a robot by using the feedback of the robot's position as detected by a vision sensor system. It is noted that in low visibility tasks, processes use RGB-D as depth cameras, where conventional cameras do not provide good estimates of sparse features of an object. In such situations, laser scanners are deemed to be the best solution in hand [1]. The stereo vision system is utilized for different robot automation applications. It has been used in a wide range of applications-from simple guidance applications to more complex systems which use data from multiple sensors. This coherent approach brings more precise objection detection and robotic manipulation in object handling. In another research, robot-assisted assembly system for installation was proposed for a variety of small components in the aircraft assembly. This system was designed and proposed to improve accuracy in the assembly process and increase the production efficiency with minimal defects on the aircraft assembly line [8]. For remote robotics, manipulators constituted with flexible camera platforms are suggested for pan-tilt units. These can be used for monitoring with active camera control and obstacle avoidance for industrial manipulators based on weighted pseudo-inverse redundancy resolution method [9]. In an extension, a new redundancy control method for robot manipulator visual servoing was proposed, which enabled obstacle detection and avoidance while end-effector motion was performed to reach the desired target position [10]. In particular, the necessary image features of the manipulator were estimated from calibrated camera parameters and the known kinematic model of the manipulator for obstacle avoidance.

In general, for any big and small sized robot, it is very difficult to visualize the object of interest as well as environmental parameters including location and orientation without using a vision system. In [11], a camera-joint coordinate mapping system was used along with a neural network for tracking-and intercept-module which determined the trajectory for a robot moving target. Moreover, an interaction control strategy was used for industrial robot manipulators which consisted of a combination of calibration free vision-based control method and impedance control approach for path tracking tasks [12]. The experiments were conducted on Fanuc M16-iB industrial robot with a force/torque sensor placed at the wrist. A camera space manipulation system which trade-offs between camera view parameters and axial deflection parameters model was proposed [13]. This resulted in re-estimation and maneuvering a very accurate placement of the robot end-effector at the target position. A robot vision system was developed for real-time object localization, detection, and tracking in three translations and three rotations [14]. A model-based approach was used without prior information of objects initial position. For automated manufacturing units, it was proposed that a system for pick-and-place tasks use randomly piled parts in a bin through measuring the 3D pose of an object by a 3D stereo vision sensor [15]. The system successfully detected the work piece to be picked using a geometric pattern matching method with respect to a 2D image with a wide field of view. Another system was developed for positioning a measured object inside the working area of a computerized measuring machine using 6-DOF industrial robot [16]. In all

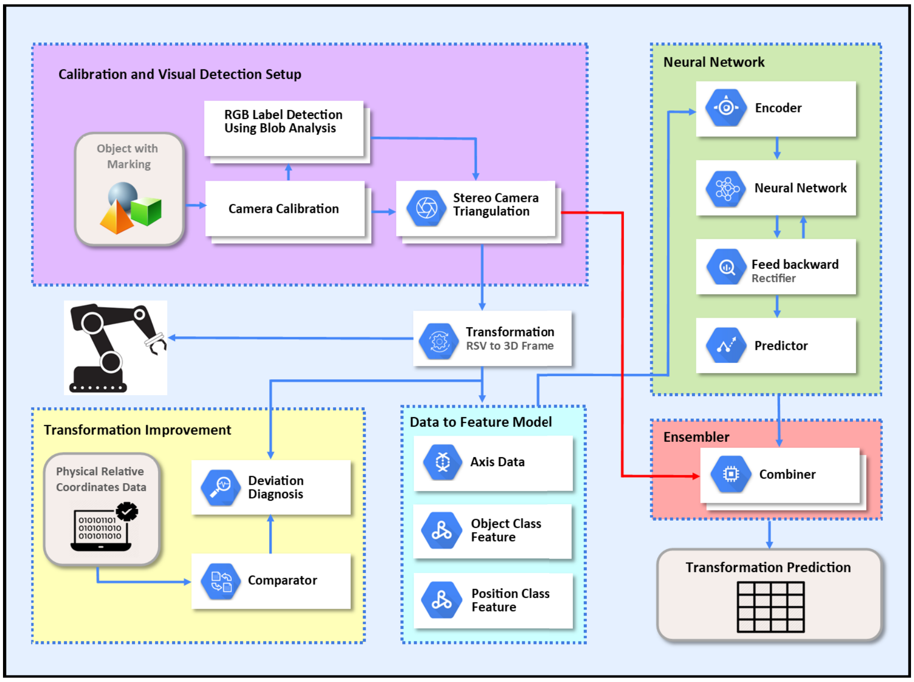


FIGURE 1. **The experimental setup used in this study to incorporate stereo vision for controlling the robotic manipulator for object detection and tracking.**

these proposed methods, there is a recent inclination towards using robotic vision system. Moreover, the performance of robotic manipulators in industrial application is known to be effected by the target object. Herein, we propose to overcome some of these challenges using a framework based on stereo vision for path-prediction and object detection.

III. MATERIALS AND METHODS

The system model diagram is shown in Figure 1 and consists of three major blocks. These include 1) visual detection using robotic stereo vision that also incorporates transformation improvement, 2) machine learning framework for path prediction using neural network, and 3) the ensemble module for robotic manipulator path prediction. Each of these modules are discussed in detail in the following text.

A. THE ROBOT STEREO VISION SYSTEM (RSV) FOR OBJECT DETECTION

Here, our aim was to detect soft objects with a color marked region. Further, the 3D centroidal point detected was converted to Cartesian coordinate system (X, Y, Z) for shortest path planning in real-time. In our proposed system, the RSV system consisted of two colored GigE TCP/IP network-based charge coupled device (CCD) cameras. The specific camera model used was FL3-GE-13S2C-CS. In particular, these are 1.3 mega pixel cameras, providing a resolution of 1288 × 964 pixels and 30 frames per second (fps). The data could be recorded at 8, 12, 16, and 24-bit resolution and came with FlyCapture software development kit (SDK). For camera based systems, calibration is a compulsory step for the estimation and extraction of 3D object geometry. In particular, we used a decision based system to enable object tracking based on detecting the marker (using red, green, and blue colors) placed on the object of interest. The flow diagram is shown in Figure 2.

In the RSV system we setup for our experiments, left and right CCD cameras were calibrated separately by considering the pinhole camera model. The calibration process started with forward projective mapping of a known geometry checkerboard pattern consisting of 256 corner points arranged in a 3D grid as 8 × 8 × 4. The checkerboard pattern was modeled by taking in to account first- and second-order approximations of the lens radial (symmetric and tangential) distortions. Moreover, camera (intrinsic and extrinsic) parameters were refined with all image points using an optimizing/minimization model which is given as,

$$S \times \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = A \times \begin{pmatrix} R & T \end{pmatrix} \times \begin{pmatrix} X' \\ Y' \\ Z' \\ 1 \end{pmatrix}, \tag{1}$$

where

$$A = \begin{bmatrix} \alpha_x & 0 & u_0 \\ 0 & \alpha_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{2}$$

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}, \tag{3}$$

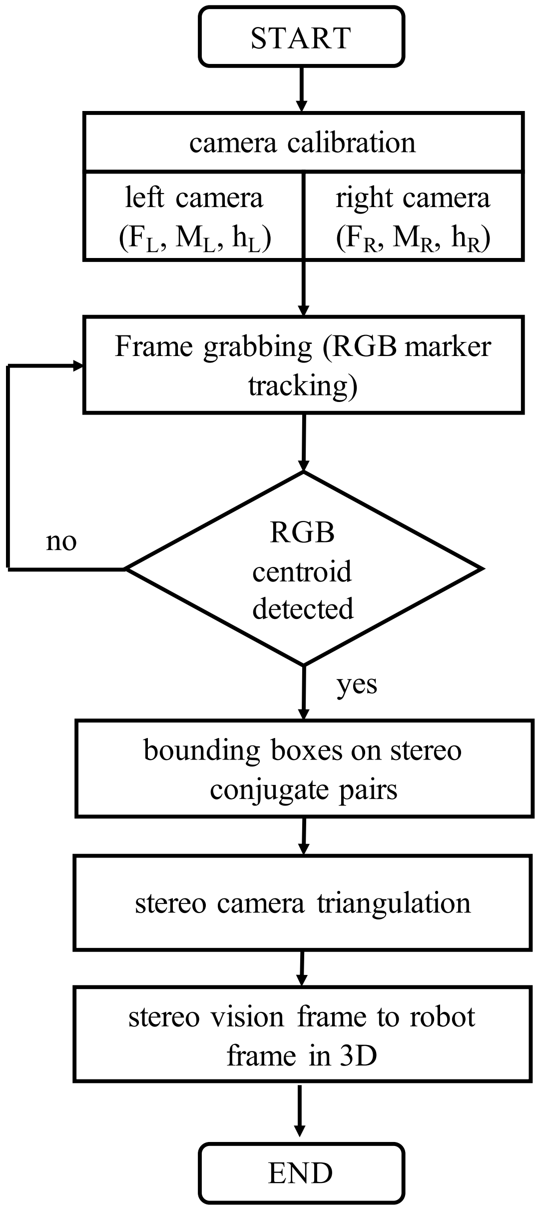


FIGURE 2. The proposed process flow of the RSV system for tracking RGB color marker in the 3D-world frame of reference.

and

$$T = \begin{bmatrix} r_{14} \\ r_{24} \\ r_{34} \end{bmatrix}. \tag{4}$$

Such that $\alpha_x = f/dx$ and $\alpha_y = f/dy$, where dx and dy are scale factors measured in mm/pixel along the x-axis and y-axis, respectively. Moreover, X, Y, and Z represents the axes in three dimensions and u_0, v_0 represent the center point for the camera in the image plane.

The camera parameters were calculated by following a two-stage camera calibration method [17]. In particular we estimated the camera homography (h), camera intrinsic (F), and camera extrinsic (M) matrices for both the left and right camera as follows

$$F_l = \begin{bmatrix} \alpha_{x1} & 0 & u_{1,0} \\ 0 & \alpha_{y1} & v_{1,0} \\ 0 & 0 & 1 \end{bmatrix}, \tag{5}$$

$$M_l = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{6}$$

$$h_l = \begin{bmatrix} i_{11} & i_{12} & i_{13} & i_{14} \\ i_{21} & i_{22} & i_{23} & i_{24} \\ i_{31} & i_{32} & i_{33} & 1 \end{bmatrix}, \tag{7}$$

$$F_r = \begin{bmatrix} \alpha_{x2} & 0 & u_{2,0} \\ 0 & \alpha_{y2} & v_{2,0} \\ 0 & 0 & 1 \end{bmatrix}, \tag{8}$$

$$M_r = \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ b_{31} & b_{32} & b_{33} & b_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{9}$$

$$h_r = \begin{bmatrix} j_{11} & j_{12} & j_{13} & j_{14} \\ j_{21} & j_{22} & j_{23} & j_{24} \\ j_{31} & j_{32} & j_{33} & 1 \end{bmatrix}, \tag{10}$$

where F_l, M_l *and* h_l are the parameters for the lest camera and $F - r, M_r,$ *and* h_r are the right camera parameters.

Since, object detection and tracking is a challenging task, there are multiple methods proposed to tackle this problem. Generally, these methods are categorized in two groups: 1) object detection based on shape and 2) object detection using special features (e.g. a marker placed on the object). Herein, we used the second approach, which uses a marker placed on the object of interest. In particular, searching an RGB marker was found to be more efficient in a constrained environment we are aiming to build. In particular, we used the blob analysis technique for the tracking task by using a color RGB marker. In general, blob analysis is a fundamental technique of machine vision based on analysis of consistent image regions. Hence, it is a tool of choice for applications in which the objects being inspected are clearly discernible from the background. In terms of implementation, we used the MATLAB computer vision toolbox to detect the object of interest (www.mathworks.com). This detection process specifically tracks regions that differ in properties like brightness or color from the rest of the image [18]. Herein, our aim was to use this method to overcome the search fatigue of a 3D object centroid point, $P(X^r, Y^r, Z^r)$, in the world frame of reference. Hence, we identified the RGB marker from the rest of the environment. The soft object having a unique color marker centroid, exists as conjugate pairs in the left and right camera image planes, represented as $a(u_1, v_1)$ and $b(u_2, v_2)$. This helps in constructing correspondence in the stereo-based vision system. We further evaluated the depth of bottle marker centroid point coordinates using single camera calibration-based homograph matrices (h_l and h_r) [19] and the pinhole camera model as

$$S_1 \times \begin{pmatrix} u_1 \\ v_1 \\ 1 \end{pmatrix} = h_l \times \begin{pmatrix} X^r \\ Y^r \\ Z^r \\ 1 \end{pmatrix}, \tag{11}$$

$$S_2 \times \begin{pmatrix} u_2 \\ v_2 \\ 1 \end{pmatrix} = h_r \times \begin{pmatrix} X^r \\ Y^r \\ Z^r \\ 1 \end{pmatrix}. \tag{12}$$

These can be solved for the X^r, Y^r, Z^r coordinates as follows

$$\begin{bmatrix} X^r \\ Y^r \\ Z^r \end{bmatrix} = (S^t S)^{-1} \times (S^t P). \tag{13}$$

Hence, we evaluated the position of RGB marked object in 3D world frame from the stereo camera using a triangulation method. Further, it was required to transform these 3D coordinates to the robotic manipulator frame coordinate system [20]. In this regard, we need to estimate ${}^RP$ (obtained from robot forward kinematics) from P (obtained from the stereo vision system) which was obtained using the triangulation method. In particular, we used the following transformation to transform from vision system to robot frame of reference

$${}^RP = ({}^RT_V)^{-1} \times P \tag{14}$$

Since we used n observations in our model to efficiently transform form the vision frame to the robot frame of reference, Equation 14 can be written as,

$${}^RP_i = ({}^RT_V)^{-1} \times P_i, \tag{15}$$

where $i = 1 \ldots n$. In particular, since we used a rigid transformation model for transformation [21], the rotation matrix (R) had to satisfy these two constraints:

$$R \times R' = I$$
$$\|R\| = 1$$

Finally, the deviation of real-time object placement was compared with coordinates of physical relative data keeping a closed experimental environment. In particular, this transformation improvement contributed towards better outcome for accuracy in positioning of the target placement of object.

B. MACHINE LEARNING FOR CONTROLLING ROBOTIC MANIPULATOR

This module consisted of feature extracting and using neural networks to learn from data. This learning mechanism was used to control the robotic manipulator for efficient path planning.

1) FEATURE EXTRACTION

Feature extraction deals with building and driving key values, called the features, from a quantifiable measured data in the field of machine learning under paradigms of pattern recognition for textual and image data. This act is solely anticipated to achieve a set of unique, informative aspect of the data in a process driven by leaning, and thus create a better interpretations with the complex data. The features used in this study represented the axes data for robotic manipulation. Since the robot frame data (X, Y, Z) is high dimensional and our machine learning problem is based on regression, we adopted non-linear principle component analysis (NLPCA) for prediction of high dimensional trajectory for displacement of robots with the features. In particular, we chose NLPCA, a special type of principal component analysis (PCA), since our framework dealt with high dimensional data and therefore a high chance of over-fitting data points. In particular, the following feature was used, where for each input value i we had a corresponding output j

$$C = \begin{bmatrix} cov(i,i) & cov(i,j) \\ cov(j,i) & cov(j,j) \end{bmatrix}, \tag{16}$$

where

$$cov(i,j) = \sum_{k=1}^{n} \frac{(i_k - \bar{i})(j_k - \bar{j})}{n-1}. \tag{17}$$

2) NEURAL NETWORK WITH NLPCA

In robotic manipulator development, the role of neural network is known to be significant [22]. One the most important factors in the automation of a robotic manipulator is the process of decision making in real-time [23]. Generally, industrial robots are controlled by using an embedded controller, however in most cases there is a clear window for significant improvements in the control system. To this end, adaptive control of robotic manipulator for both linear and non-linear systems has seen a significant interest [24], [25]. This interest is largely due to the fact that adaptive control theory is particularly well-suited to robot manipulators where the dynamic model is highly complex and may contain several unknown parameters. For output optimization and an increase in system reliability, neural networks are well suited and range from a single DOF to multi DOF applications. From programming a simulated 2D motion robotic arm to sense the environment variables, variants of machine learning (ML) algorithms like convolutional neural networks (CNN) and recursive-CNN are widely used [26].

Since the transformation data of robot is high dimensional and non linear, we propose to use non-linear principle component analysis with a neural network and call this NLPCA-NN. The framework is represented as an auto encoder with the use of a set of cascaded multi-layer NN with a bottleneck structure to extract nonlinear information of the data set. In particular, to understand the mechanism of the proposed NLPCA with neural network we consider a function f, which represents the transformation $R_p \rightarrow R_r$. This part of

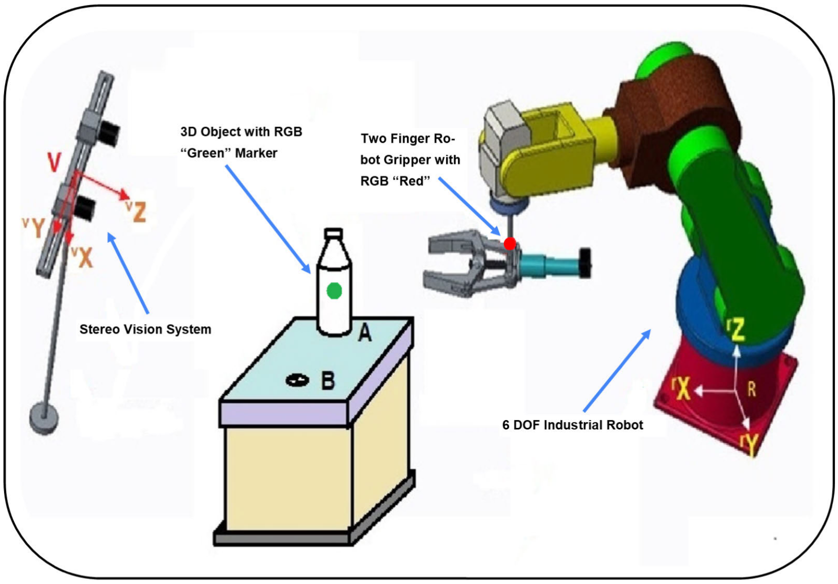


FIGURE 3. The experimental setup used in this study consisting of a bottle, stereo vision cameras and a robotic manipulator.

transformation was represented by the first three layers of our proposed NN, where the first layer was the input layer and the third layer was the bottleneck layer. Suppose another function s such that $s : R_r \rightarrow R'_p$. This was performed by the fourth and fifth layers of the proposed neural network, where the fifth layer was the output layer. Using this notation, the weights in the NLPCA network were determined using the following objective function

$$min \sum_{i=1}^{n} \left\| {}^{\rho}X_i - {}^{\rho} X'_i \right\| \tag{18}$$

where ${}^{\rho}X'$ is the output of the network. In general f can be any non-linear function, which in our proposed framework is the feed-forward NN which maps the input data to the bottleneck layer. For the proposed neural network, the weights were randomly initialized. The hyper-parameters, including the learning rate, were selected using grid search. For training we used 10-fold cross validation, where data were split in ten equal parts and each part was used for testing while the other nine parts were used for training. This process was repeated for all ten parts and the average results were reported. The neural network was a feed-forward neural network and the training was performed using stochastic gradient descent. In particular, we used 10 epochs and observed that in all experiments, the training error stabilized after these epochs, hence a higher number was not selected.

C. ENSEMBLER

The proposed framework coupled with the neural network increased robot accuracy for completion of the pick and place task for a soft object in real time. This was enabled by combining calibration and visual detection step with the neural network based predictor. Herein, we call this step as the ensembler. The role of ensembler was to effectively increase the precise nature of object displacement by considering both the object RGB characteristics and its features in a joint manner.

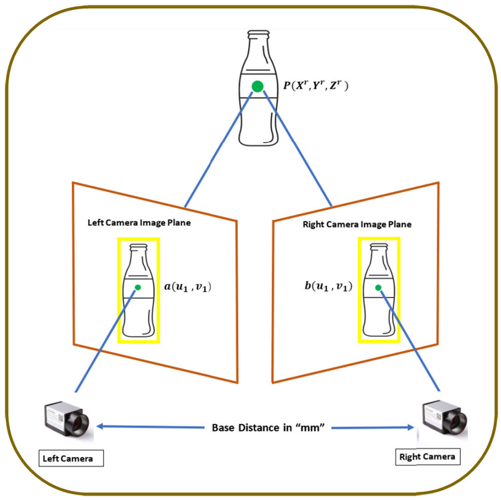
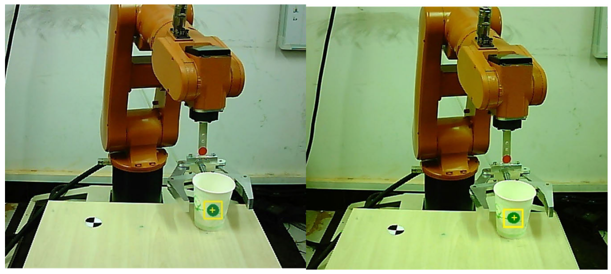


FIGURE 4. The stereo vision frame of reference used for detecting the RGB marker placed on the soft object.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The experimental setup of our proposed framework is shown in Figure 3. The model includes un-rectified parallel dual stereo cameras with an object in the middle. Further, a 6-DOF robotic manipulator and two-finger gripper was also attached. Our proposed design was developed considering an RSV system to avoid extra computational requirements and reduce errors in object detection. In particular, we designed an RGB

(a) An example of RGB marker detection using the proposed technique on a soft object.



(b) An example of RGB marker tracking in real-time using video frames.

FIGURE 5. **RGB marker tracking using the proposed RSV framework.**

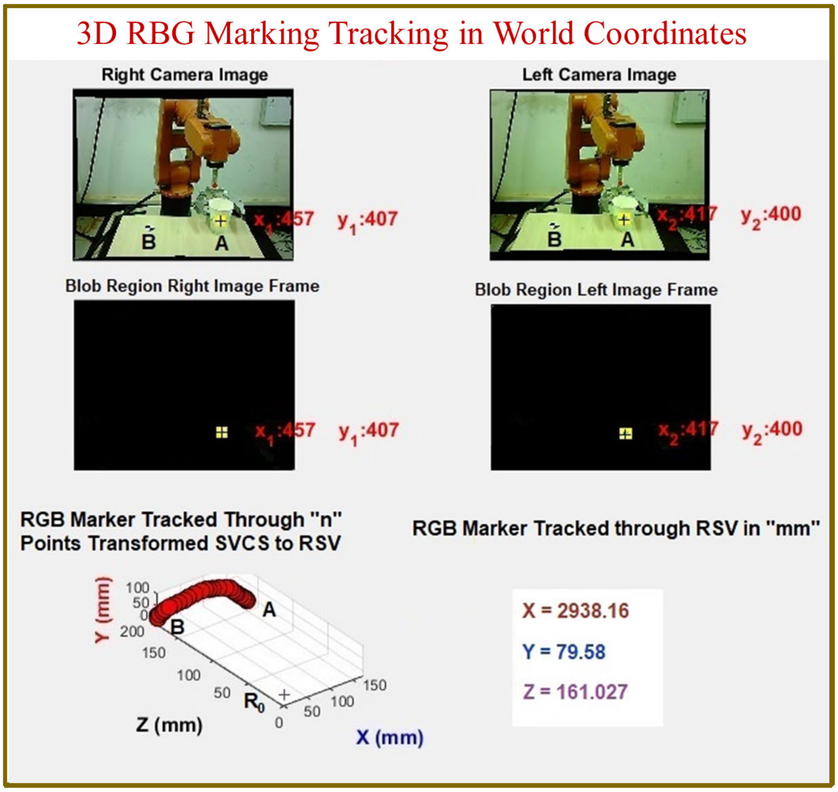


FIGURE 6. **A graphical user interface to observe errors in frame of reference transformation.**

tracker, which tracks the RGB marker placed on the target object, to track the object in 3D with feature extraction in real-time. We also designed a module for the robotic manipulator using an ensemble technique, enabling an efficient computation of target trajectories.

A. OBJECT TRACKING USING RGB MARKER

The RSV system used three colors including red, green, and blue coloring system to track the objects which were duly marked. The RVS system has to be calibrated first, by knowing transformation from robot frame points in 3D world frame before execution of our desired tasks for the unknown 3D position of RGB marked objects in RVS system. From single camera, calibration parameters were obtained for the left and right cameras. Moreover, blob analysis and triangulation techniques were used to detect the marker. The information form the left and right image planes was combined to generate the coordinates of the marker placed on the object. A representation of this framework is presented in Figure 4.

In our experiments we setup soft objects including bottles and cups and placed RGB markers. We experimented with the RSV system to ensure that the marker is detected as well as tracked in real-time from videos. A couple of examples are shown in Figure 5, where RGB markers were placed on the soft objects and the marker position was calculated in the 3D

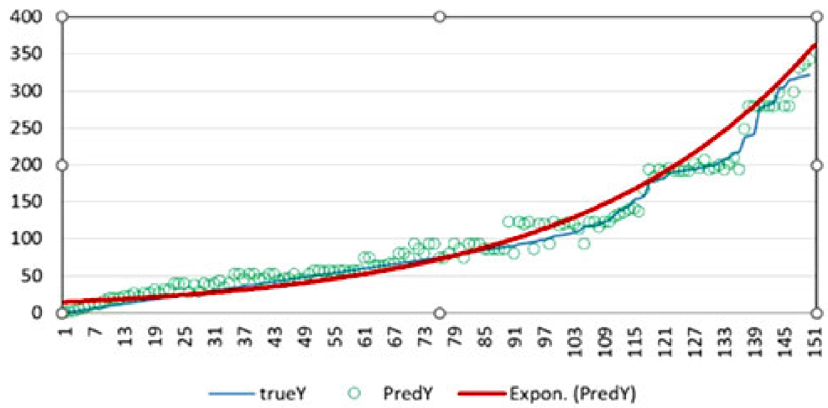


FIGURE 7. A comparative plot representing the predicted points and the actual points.

TABLE 1. A summary of experiments performed on various soft objects and the corresponding error rates.

Object Class	Object Weight (grams)	Iterations	Variable	Error Rate (Robot Frame Vs Stereo Vision Frame) Mean	Median
Filled	307	120	X	3.01e-12	-0.84
Filled	307	120	Y	-2.12e-13	-0.02
Filled	307	120	Z	7.76e-13	-1.69
Semi-filled	136	133	X	-0.02	0.15
Semi-filled	136	133	Y	0.01	-0.27
Semi-filled	136	133	Z	0.50	0.62
Empty	7	118	X	-0.02	-0.13
Empty	7	118	Y	0.01	0.17
Empty	7	118	Z	-0.65	-1.24

world frame of reference. Further, parametric transformations were applied to transform from RVS system to robot frame of reference. This was required to calculate how much robot end-effector has to move in the X, Y, and Z axes.

Due to generation of misaligned information from the stereo cameras some hardware inaccuracies were recorded. Hence, multiple iterations were performed for RGB marker tracking and trajectory generation. For evaluating these results a graphical user interface was designed, which displayed the error in trajectory calculation in the 3D frame of reference.

B. TRAJECTORY PLANNING

In general, robot manipulators principally operate with both dependent and independent joint controllers. The independent controller controls the joint angles distinctly via simple constant gain by means of position servo loops. This allows the manipulator to perform ordinary positioning like placement and picking operations to positions called the *object displacement*. However, because of separate controller to each section, it is ultimately limited in precise tracking performance. Such tracking creates dedicate trajectories [27], to the target place for placement of an object. For active trajectories and obtaining sustainable manipulator performance, the execution of the robot is dependent on objects and their associated payloads. Herein, we were particularly interested in soft objects and used machine learning for better learning the trajectories. The transformation prediction module performed path planning after ensembling both the RSV and NN predictions. The results are shown in Figure 7, where $true_Y$ represents the real points of the object to place and $Pred_Y$ is the predicted points using the proposed framework. Further, the error value is also plotted and it can be observed that a significant performance was achieved in path tracking with minimal error.

A comprehensive summary of the experiments performed is presented in Table 1. We considered various parameters in our experiments. In particular, experiments were performed with various categories of soft objects including filled, semi-filled, and empty. This change in payload is known to effect the trajectory planning and effort required to pick the object. We also recorded the actual weight of the object and is reported in the table. Further, we used various number of iterations for each object category and report a mean and median error rate between the robot frame and the stereo vision frame of reference. In particular, these error values are significantly small and hence our proposed framework was highly effective in this transformation. It should be noted that with this performance, the proposed framework can benefit the utilization of robotic manipulators in industrial applications, particularly targeted towards handling soft objects. Moreover, we calculated the accuracy in path planning using the error value in predicting the trajectory points. In our experiments, we achieved an accuracy of 91.8%, which is significant when dealing with soft objects. Hence, our proposed framework benefits from stereo vision and machine learning to enable better object tracking and trajectory planning when used with robotic manipulators.

V. CONCLUSION

Herein, we solve the path planning and object positioning problem using robot stereo vision system. The proposed system is assisted by a neural network-based method followed by an ensemble framework. The path planning frameworks have benefited from dual camera positioning, thereby improving both localization as well as better outcomes in the production line in various industries. For instance in the bottling industry, the nature of object involved is delicate, hence precision remains the most significant task. To this end, we developed a robot vision experiment platform for a 6-DOF industrial robot, which was highly accurate in detecting objects of interest. After detection of object (using an RGB marker), we used an object placement method, called the transformation, which generates the data points for object placement in real-time. To this end, we used a feature extraction module- non-linear principal component analysis- along with a neural network for trajectory prediction. We had to combine the RSV outcomes with robot visual system for more accurate transformation predictions. Our experimental results show that the proposed framework can effectively improve the accuracy of path planning and enhance object detection via RSV. Future work would include identifying objects, tracking path, and distinguishing object types with more dynamic nature of targeted objects. A potential effort could be towards speeding-up the overall system processing time to achieve high throughput in industrial robotics.

REFERENCES

[1] R. C. Leishman, D. P. Koch, T. W. McLain, and R. W. Beard, "Robust visual motion estimation using RGB-D cameras," in *Proc. AIAA Infotech Aerosp. Conf.*, 2013, pp. 1–13.

[2] D. Guo, Z. Li, A. H. Khan, Q. Feng, and J. Cai, "Repetitive motion planning of robotic manipulators with guaranteed precision," *IEEE Trans. Ind. Informat.*, vol. 17, no. 1, pp. 356–366, Jan. 2020.

[3] F. Zha, Y. Fu, P. Wang, W. Guo, M. Li, X. Wang, and H. Cai, "Semantic 3D reconstruction for robotic manipulators with an eye-in-hand vision system," *Appl. Sci.*, vol. 10, no. 3, p. 1183, Feb. 2020.

[4] C. Mineo, M. Vasilev, B. Cowan, C. N. MacLeod, S. G. Pierce, C. Wong, E. Yang, R. Fuentes, and E. J. Cross, "Enabling robotic adaptive behaviour capabilities for new industry 4.0 automated quality inspection paradigms," *Insight Non-Destructive Test. Condition Monitor.*, vol. 62, no. 6, pp. 338–344, Jun. 2020.

[5] S. Baressi Šegota, N. Anelić, I. Lorencin, M. Saga, and Z. Car, "Path planning optimization of six-degree-of-freedom robotic manipulators using evolutionary algorithms," *Int. J. Adv. Robotic Syst.*, vol. 17, no. 2, 2020, Art. no. 1729881420908076.

[6] S. Wang, X. Shao, L. Yang, and N. Liu, "Deep learning aided dynamic parameter identification of 6-DOF robot manipulators," *IEEE Access*, vol. 8, pp. 138102–138116, 2020.

[7] J. Mattila, J. Koivumaki, D. G. Caldwell, and C. Semini, "A survey on control of hydraulic robotic manipulators with projection to future trends," *IEEE/ASME Trans. Mechatronics*, vol. 22, no. 2, pp. 669–680, Apr. 2017.

[8] Z. Jin, C. Yu, J. Li, and Y. Ke, "A robot assisted assembly system for small components in aircraft assembly," *Ind. Robot, An Int. J.*, vol. 41, no. 5, pp. 413–420, Aug. 2014.

[9] M. Bjerkeng, A. A. Transeth, K. Y. Pettersen, E. Kyrkjebo, and S. A. Fjerdingen, "Active camera control with obstacle avoidance for remote operations with industrial manipulators: Implementation and experimental results," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Sep. 2011, pp. 247–254.

[10] M. Mikawa, K. Yoshida, M. Tanno, and M. Matsumoto, "Vision-based redundancy control of robot manipulators for obstacle avoidance," in *Proc. 23rd Int. Conf. Ind. Electron., Control, Instrum.*, 1997, pp. 1373–1378.

[11] Y. Chen and C. Hsueh, "Complementary data fusion in vision-guide and control of robotic tracking," *Robotica*, vol. 19, no. 1, p. 53, 2001.

[12] I. Bonilla, M. Mendoza, E. J. Gonzalez-Galván, C. Chavez-Olivares, A. Loredo-Flores, and F. Reyes, "Path-tracking maneuvers with industrial robot manipulators using uncalibrated vision and impedance control," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 42, no. 6, pp. 1716–1729, Nov. 2012.

[13] G. Parker, and E. T. Baumgartner, "Robot positioning of flexible-link manipulator using vision," *Robotica*, vol. 22, no. 3, pp. 301–307, May 2004.

[14] Y. Sumi, Y. Ishiyama, and F. Tomita, "Robot-vision architecture for real-time 6-DOF object localization," *Comput. Vis. Image Understand.*, vol. 105, no. 3, pp. 218–230, Mar. 2007.

[15] J.-K. Oh, S. Lee, and C.-H. Lee, "Stereo vision based automation for a bin-picking solution," *Int. J. Control, Autom. Syst.*, vol. 10, no. 2, pp. 362–373, Apr. 2012.

[16] S. Lemes, D. Strbac, and M. Cabaravdic, "Using industrial robots to manipulate the measured object in CMM," *Int. J. Adv. Robotic Syst.*, vol. 10, no. 7, p. 281, Jul. 2013.

[17] C. Jie, "Two stage camera calibration modeling and simulation," in *Proc. 6th Int. Bhurban Conf. Appl. Sci. Technol.*, Dec. 2009, pp. 231–237.

[18] T. Mahalingam and M. Subramoniam, "A robust single and multiple moving object detection, tracking and classification," *Appl. Comput. Informat.*, Jul. 2020.

[19] M. Xie, *Fundamentals of Robotics: Linking Perception to Action*, vol. 54. Singapore: World Scientific, 2003.

[20] A. Shahzad, Y. Mu, L. Zhang, D. Cui, Y. L. H. Dong, and X. Gao, "Simulation and modelling of 6-degree of freedom STEP SD 700 industrial robot follow-up on a given trajectory," in *Proc. Chin. Control Decis. Conf. (CCDC)*, May 2016, pp. 6751–6755.

[21] Z. Min, J. Wang, and M. Q.-H. Meng, "Joint rigid registration of multiple generalized point sets with hybrid mixture models," *IEEE Trans. Autom. Sci. Eng.*, vol. 17, no. 1, pp. 334–347, Jan. 2020.

[22] F. Luan, J. Na, Y. Huang, and G. Gao, "Adaptive neural network control for robotic manipulators with guaranteed finite-time convergence," *Neurocomputing*, vol. 337, pp. 153–164, Apr. 2019.

[23] B. Rahmani and M. Belkheiri, "Adaptive neural network output feedback control for flexible multi-link robotic manipulators," *Int. J. Control*, vol. 92, no. 10, pp. 2324–2338, Oct. 2019.

[24] Q. Zhou, S. Zhao, H. Li, R. Lu, and C. Wu, "Adaptive neural network tracking control for robotic manipulators with dead zone," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 12, pp. 3611–3620, Dec. 2019.

[25] D. Zhang and B. Wei, "A review on model reference adaptive control of robotic manipulators," *Annu. Rev. Control*, vol. 43, pp. 188–198, Dec. 2017.

[26] S. Wan and S. Goudos, "Faster R-CNN for multi-class fruit detection using a robotic vision system," *Comput. Netw.*, vol. 168, Feb. 2020, Art. no. 107036.

[27] Z. Feng and Y. Zhu, "A survey on trajectory data mining: Techniques and applications," *IEEE Access*, vol. 4, pp. 2056–2067, 2016.

ADIL SHAHZAD was born in Peshawar, Khyber Pakhtunkhwa, Pakistan, in 1973. He received the B.E. degree in mechanical engineering from the Balochistan University of Engineering and Technology Khuzdar (BUETK), in 1998, and the M.S. degree in automation and manufacturing engineering from the School of Vehicular Engineering, Beijing Institute of Technology (BIT), Beijing, China, in 2009, where he is currently pursuing the Ph.D. degree with the School of Mechatronical Engineering.

XUESHAN GAO (Member, IEEE) graduated from the School of Mechatronics Engineering, Harbin Institute of Technology, China. He received the M.S. degree from the Faculty of Engineering, University of Miyazaki, Japan, in 1996, and the Ph.D. degree from the School of Electrical Engineering and Automation, Harbin Institute of Technology, in 2002. From 2003 to 2005, he was a Guest Professor with the Chiba Institute of Technology, Japan. Since 2005, he has been with the Intelligent Robotics Institute, Beijing Institute of Technology, China. His research interests include mobile robot, industrial robot, and evolutionary robot.

AWAIS YASIN received the B.S. degree from the University of Engineering and Technology, Taxila, and the M.S. and Ph.D. degrees in the field of robotics from the Beijing Institute of Technology, Beijing, China, in 2013. He has a rich industrial experience of more than 16 years serving different research organizations of Pakistan. He is currently the Director of the Virtual Reality and Machine Learning Laboratory, UET Taxila, an Associate Professor with NUTECH, Islamabad, and a member of Industrial Advisory Board (IAB) for more than ten different universities of Pakistan. His research and teaching interests include robotics, machine learning, virtual reality, drive simulation, and intelligent transportation systems. He was honored to win the World Cup Award in the field VR Cloud Programming held in Tokyo, Japan, in 2018.

KAMRAN JAVED received the Ph.D. degree in automatic control engineering. He is currently serving as an Associate Professor and the Head of the Computer Engineering Department, NUTECH, Islamabad, Pakistan. His research interests include machine learning, computational intelligence, and robotics.

SYED MUHAMMAD ANWAR (Member, IEEE) received the B.Sc. degree (Hons.) in computer engineering from the University of Engineering and Technology (UET) Taxila, Pakistan, in 2005, and the M.Sc. degree (Hons.) in data communications and the Ph.D. degree in electronic and electrical engineering from The University of Sheffield, U.K., in 2007 and 2012, respectively. He is currently an Associate Professor with the Department of Software Engineering, UET Taxila. His research interests include medical imaging, data communication, machine learning, robotics, and human–computer interaction. He was a recipient of the prestigious Fulbright Fellowship.