INTERNATIONAL JOURNAL OF ROBOTICS
(THEORY AND APPLICATIONS)
E-mail: IJR@kntu.ac.ir

Eye In-hand Stereo Image Based Visual Servoing for Robotic Assembly and Set-Point Calibration used on 4 DOF SCARA Robot

Mahmoud Jeddi [a], Ahmad Reza Khoogar [b,*], Ali Mehdipoor Omrani [c]

[a] Ph.D. Student, Malek Ashtar University of Technology, Tehran , Iran
[b] Associate Professor, Malek Ashtar University of Technology, Tehran, Iran
[c] Associate Professor, Malek Ashtar University of Technology, Tehran, Iran

ARTICLE INFO

Article history:
Submit: 2020-11-27
Revise: 2021-06-21
Accept: 2022-12-31

Keywords:
Stereo Vision,
SCARA Robot,
Feature extraction,
Image Based Visual Servoing
Extended Kalman Filter

ABSTRACT

The method is presented for object assembling via manipulator robot. This method was designed to track the pieces based on image feedback for pick and placement task. The depth of the pieces is calculated by stereo triangulation. Vision-guided robotic is most often based on the training steps, therefore it is a time-consuming process. For this reason, we have proposed a modified stereo vision method to predict the movement of the part in the image space. The linear and angular velocities of moving objects predict by a state estimator. The object velocity components predicted by estimation algorithms such as Kalman filter, Recursive least square and Extended Kalman Filter. Results show that in the case of, Extended Kalman filter estimator shows better tracking and convergence behavior. This method does not need to know the three-dimensional model of the parts, and it can be used on pick and place robots if the physical limitations of the joints are considered. The proposed method was experimentally tested in a laboratory environment. The cameras were installed parallel and non-parallel to determine the effect of the field of view on the precision and speed detection. A comparison of the simulation and experimental results showed that the use of the parallel stereo Image-based visual servoing with Extended Kalman Filter method could be smoother and more accurate than the other methods.

[*] Corresponding address: Faculty of Material and Manufacturing Technologies, Malek Ashtar University of Technology,
Tel.: +98-21-22987686, E-mail: khoogar@mut.ac.ir

1. Introduction

The stereo vision was designed based on the concept of how humans see an object in the world and how can sense the difference depth between objects. When someone looks at the object, the left sight of the object differs from the right one. By staring at the target the two eyes converge so that the object appears at the center of the nerve layer that lines the back of the eye, senses light, and creates the impulses. The created impulse travels through the optic nerve to the brain which is called the retina. The object emerges on the center of the retina in both eyes. The third dimension, depth has been extracted from computing retinal disparity. Appending this capability to the industrial robots could enable them to autonomously accomplish manipulation tasks such as drilling, pick & placement and part assembly. Object recognition and pose detection have great importance in robotics since they help robots to localize objects [2]. In robot vision, the speed of the assembly process is strongly influenced by the visual system. The visual servoing consists of several stages, the first step is camera calibration, which defines the internal and external parameters of the camera [3], triangulation is the second step which determines the depth measurement. This info specifies the location and position of the target in space [3]. To identify and classify parts, the features of the parts are selected by the image processing system. Selected features should be processed and sent to the robot controller for decision-making and grasping components [4]. Variation in the robot's joint values occurs by the vision-based controller according to the observation and motion algorithm. Therefore, the system error is the sum of the errors caused by the motion control and the visual system. In robotic systems, where the control process is provided by visual feedback, the insufficient stability is significant due to the limitations of the cameras in achieving high image sampling rates [4]. Due to the fact that in a monocular visual servoing, the distance of the object can't be obtained instantaneously, so the geometric model of the object, as well as the distance as primary data, is needed to calculate the jacobian matrix. So the convergence rate in these systems is slowed down to calculate the control law. Stereo vision systems use the principles of epipolar geometry derived from two images, so they can detect object distance and provide faster performance characteristics without the need of predefined object data. [5]. Using the features obtained from the image of the target object, control of the situation and position of the robot end-effector relative to the target is performed [6]. Position-based and image-based visual servoing are the main types of approach in visual servoing control. Image feature is a nonlinear function of the camera pose, so the Image-based visual servoing control approach is highly challenging [7]. The relationships between the time variation of features in pixel coordinate and the camera velocity are needed to design IBVS (i.e. "Image-based visual serving") approach. In the following, a review of previous studies has been done and the advantages and disadvantages of each method have been discussed. In the third part, the research methodology is presented and the simulation results are given. In the fourth section, the obtained results are discussed and operational approaches for using this method are presented. Finally, in the conclusion section, the project findings are discussed.

1. Literature Review

Automatic assembling of sensitive and delicate components requires robust precision. In robotic assembling location and orientation of the piece should be determined concerning the robot gripper. Many results in the past years have been devoted to the visual servoing problem ([8],[9],[10],[11],[12],[13]). The PBVS (i.e. "Point-based visual serving") and IBVS methods are different in the natures, due to the inputs used in their control schemes [15]. Although the robot can reach the target in both approaches, the robot's movement behavior is not the same [14]. The divergence and instability problems may occur in some cases when the target is away from the initial position [16]. In the PBVS approach, the position of the object should be estimated so the processing speed of the IBVS approach is higher than the PBVS one. In position based method, identifying the observed image plane features and the camera's intrinsic and extrinsic parameters are necessary while in the Image-based visual servo, the pose estimation step doesn't needed because it uses directly the image features for control [4]. Various techniques have been proposed to accost the accurate depth extraction by using the stereo-vision system. In unstructured environments, stereo visual servoing has shown to be very effective compare to all other visual servoing categories. The special relationship in dual-camera system such as the epipolar line of stereo vision, and the space intersection method is essential and widely researched [17]. Rahardja and Kosaka [7] developed an algorithm to find visual features of complex objects by stereo vision. The problem of their system was needing to human intervention for grasping task. Hema [18] et al. developed an intelligent bin-picking method using stereo vision. This system was able to detect the position and distance of the objects height. The stereo system calculated distance-related information using neural network training. Their method was based on eye to hand PBVS method. Due to the fact that sometimes the robot arm is placed in front of the cameras and the view of the cameras is blocked, so the controller may fail and the robot may not be able to do its task. Also, due to the position estimation phase, this method is time consuming and

in practice has little application in fast assemblies. The mapping between the 2-D pixel coordinate and 3-D world coordinate is called camera calibration. Several classical methods developed based on the pinhole model, such as the methods developed by Tsai [19], Heikkila [20], etc. for camera calibration. These instructions also contain the necessary information for manual editing. Shaoyan Gai et al [21], proposed a new method for stereo image calibration. The calibration plate includes some regular dot points with black color while these points have a certain distance from each other in the direction of x and y. A specified number of images were taken from different viewing angles relative to the calibration plate in the measurement range. The camera calibration helps to achieve the initial parameters of cameras. So using the centroid distance matrix, the rotation and translation matrix of the cameras would be computed. For simplification, the coupling degree parameters of both cameras were reduced. Due to the space intersection method, three-dimensional coordinates of calibrated points are calculated. So the calibration parameters are optimized by minimizing the total errors. Two co-planar images plane simulate via rectification method which method rectifies the rotation of the camera hence these planes are parallel to the baseline [22]. The robot has a series of limitations on the speed and path to reach the goal because of the physical construction of the robot joints. If the constraints in visual servoing are not well defined, system can be fail. The possessed in common bug of the vision control frameworks reported in [23], where the limits of the constraints of joint-velocity and joint-displacement have not been well considered. Manipulator robots have made of a series of joints that have physical limitations. Unfortunately, the robot task may be failed and even robot damaged, if the joint limits violated during a task process [14]. In the Visual Servoing approach, the robot performs its task by using feedback from the location and position of the workpiece. Therefore, online estimating the position and condition of the workpiece is an important issue in Visual Servoing. Due to the time-consuming process of vision and decision making, so new approaches have been proposed in this field to be able to estimate the position of the parts in the next steps according to the current and previous situation. The various strategy implemented by the researchers to overcome the delay problem for predicting the trajectory of the moving object which is being manipulated. For example, an extended Kalman filter is used in vision-based pose-estimation schemes in robot control. Considering a series of initial assumptions helps the estimator technique to provide an appropriate answer. Known initial object poses in 3-D space, noise statistics, and sufficiently high sampling rates of frames, the prediction technique obtain satisfactory results. [24]. The Kalman filter was introduced to track the trajectory of spacecraft at NASA. This method was applied to predict the state of a system with noisy inputs and estimate the true current values [25]. The Kalman filter can be used as a predictor in visual servoing to estimate the state variables based on input images because of the state transition irregular components [26]. Vision system using visual features extracted from an image of the goal object control the pose of the robot's end-effector, relative to the object. For prediction the moving object, proposed the new method based on recursive least square and Broyden method to partitioned the Jacobian matrix [27]. This paper goal is to optimize the assembly of robotic parts without the need for an operator. To increase the accuracy of assembling the two cameras are installed on the robot's end-effector to detect the position and direction of the parts. Finally, the processed information is sent to the robot controller for decision. This article is written in several sections. First in section 3.1 the dynamical model of IBVS is introduced then at section 3.2 Interaction Matrix for Stereo vision are described. Finally, the controller is designed based on error reduction. In section 4, the simulation and experimental results are discussed. At the end, the advantage of the presented method is explained.

2. Dynamical model of the Stereo-IBVS

By mounting the camera on the robot end-effector, we will have in the world frame $\vartheta = (v, \omega)$ as a body velocity and $P = (X, Y, Z)$ as a world point in camera relative coordinate. The point velocity can be described in the camera frame as,

$$= -\boldsymbol{\omega} \times \boldsymbol{P} - \boldsymbol{v} \quad (1$$

This equation could be write in scalar form as,

$$\dot{\boldsymbol{X}} = \boldsymbol{Y\omega_X} - \boldsymbol{Z\omega_Y} - \boldsymbol{v_X}, \dot{\boldsymbol{Y}} = \boldsymbol{Z\omega_X} - \boldsymbol{X\omega_z} - \boldsymbol{v_Y}, \dot{\boldsymbol{Z}} = \boldsymbol{X\omega_Y} - \boldsymbol{Y\omega_X} - \boldsymbol{v_Z} \quad (2$$

by normalized image-plane, the perspective projection coordinate could write as,

$$\boldsymbol{x} = \frac{\boldsymbol{X}}{\boldsymbol{Z}}\ , \boldsymbol{y} = \frac{\boldsymbol{Y}}{\boldsymbol{Z}} \quad (3$$

using the quotient rule drive the temporal derivative as,

$$\dot{\boldsymbol{x}} = \frac{\dot{\boldsymbol{X}}\boldsymbol{Z} - \boldsymbol{X}\dot{\boldsymbol{Z}}}{\boldsymbol{Z^2}}\ , \dot{\boldsymbol{y}} = \frac{\dot{\boldsymbol{Y}}\boldsymbol{Z} - \boldsymbol{Y}\dot{\boldsymbol{Z}}}{\boldsymbol{Z^2}} \quad (4$$

with placement $X = xZ$ and $Y = yZ$in Eq.2 then write in matrix form as,

$$\begin{pmatrix}\dot{\boldsymbol{x}}\\ \dot{\boldsymbol{y}}\end{pmatrix} = \begin{bmatrix} -\frac{\boldsymbol{1}}{\boldsymbol{Z}} & \boldsymbol{0} & \frac{\boldsymbol{x}}{\boldsymbol{Z}} & \boldsymbol{xy} & -(\boldsymbol{1}+\boldsymbol{x^2}) & \boldsymbol{y} \\ \boldsymbol{0} & -\frac{\boldsymbol{1}}{\boldsymbol{Z}} & \frac{\boldsymbol{y}}{\boldsymbol{Z}} & \boldsymbol{1}+\boldsymbol{y^2} & -\boldsymbol{xy} & -\boldsymbol{x} \end{bmatrix} \begin{bmatrix} \boldsymbol{v_x} \\ \boldsymbol{v_y} \\ \boldsymbol{v_z} \\ \boldsymbol{\omega_x} \\ \boldsymbol{\omega_y} \\ \boldsymbol{\omega_z} \end{bmatrix} \quad (5$$

The camera spatial velocity mapped to feature velocity at normalized image coordinate. So mapping image-plane coordinate to pixel coordinate could be write,

$$\boldsymbol{u}=\frac{f}{\rho_u}\boldsymbol{x}+\boldsymbol{u_0}\ ,\boldsymbol{v}=\frac{f}{\rho_v}\boldsymbol{y}+\boldsymbol{v_0} \quad (6$$

In Eq.6 f describe the focus length and(u_0, v_0) is principle point. By definition $\bar{u}=u-u_0$ and $\bar{v}=v-v_0$ Eq.6 could be rearranged as $x=\frac{\rho_u}{f}\bar{u}$, $y=\frac{\rho_v}{f}\bar{v}$ by considering ρ_u , $\rho_v$and $f$is constant parameters, so the temporal derivative related to the pixel coordinates is,

$$\boldsymbol{x}=\frac{\rho_u}{f}\dot{\bar{\boldsymbol{u}}}\ ,\boldsymbol{y}=\frac{\rho_v}{f}\dot{\bar{\boldsymbol{v}}} \quad (7$$

and substituting Eq.7 and Eq.5 into Eq.6 leads to,

$$\begin{pmatrix}\dot{\bar{\boldsymbol{u}}}\\ \dot{\bar{\boldsymbol{v}}}\end{pmatrix}=\begin{bmatrix}-\frac{f}{\rho_u Z} & \mathbf{0} & \frac{\bar{u}}{Z} & \frac{\rho_v\bar{u}\bar{v}}{f} & -\frac{f^2+\rho_u{}^2\bar{u}^2}{\rho_u f} & \frac{\rho_v\bar{v}}{\rho_u}\\ \mathbf{0} & -\frac{f}{\rho_v Z} & \frac{\bar{v}}{Z} & \frac{f^2+\rho_v{}^2\bar{v}^2}{\rho_v f} & -\frac{\rho_v\bar{u}\bar{v}}{f} & -\frac{\rho_u\bar{v}}{\rho_v}\end{bmatrix}\begin{bmatrix}\boldsymbol{v_x}\\ \boldsymbol{v_y}\\ \boldsymbol{v_z}\\ \boldsymbol{\omega_x}\\ \boldsymbol{\omega_y}\\ \boldsymbol{\omega_z}\end{bmatrix} \quad (8$$

to simplify assume $\rho=\rho_u=\rho_v$ and $\bar{f}={}^{f}/_{\rho}$. The Jacobian matrix could be simplified as,

$$\boldsymbol{J_p(p,Z)}=\begin{bmatrix}-\frac{\bar{f}}{Z} & \mathbf{0} & \frac{\bar{u}}{Z} & \frac{\bar{u}\bar{v}}{\bar{f}} & -\frac{\bar{f}^2+\bar{u}^2}{f} & \bar{v}\\ \mathbf{0} & -\frac{\bar{f}}{Z} & \frac{\bar{v}}{Z} & \frac{\bar{f}^2+\bar{v}^2}{f} & \frac{\bar{u}\bar{v}}{\bar{f}} & -\bar{u}\end{bmatrix} \quad (9$$

Eq.7 can write in short form matrix as,

$$\dot{\boldsymbol{P}}=\boldsymbol{J_p(p,Z)\vartheta} \quad (10$$

where $J_p$is the image Jacobian matrix with two rows and six column that is represent a camera distance Z and point feature at coordinate p. Jacobian matrix $J_p(p,Z)$ could be divided in two parts of translation and angular,

$$\boldsymbol{J_p(p,Z)}=\begin{bmatrix}\underbrace{\begin{matrix}-\frac{\bar{f}}{Z} & \mathbf{0} & \frac{\bar{u}}{Z}\\ \mathbf{0} & -\frac{\bar{f}}{Z} & \frac{\bar{v}}{Z}\end{matrix}}_{\frac{1}{Z}J_t} & \underbrace{\begin{matrix}\frac{\bar{u}\bar{v}}{\bar{f}} & -\frac{\bar{f}^2+\bar{u}^2}{f} & \bar{v}\\ \frac{\bar{f}^2+\bar{v}^2}{f} & \frac{\bar{u}\bar{v}}{\bar{f}} & -\bar{u}\end{matrix}}_{J_\omega}\end{bmatrix} \quad (11$$

the Eq.11 can write in brevity as,

$$\dot{\boldsymbol{P}}=\left(\frac{1}{Z}\boldsymbol{J_t(p,Z)}\ \vdots\ \boldsymbol{J_\omega(p,Z)}\right)\boldsymbol{\vartheta} \quad (12$$

substitute Eq. 12 into Eq. 8,

$$\begin{pmatrix}\dot{\bar{\boldsymbol{u}}}\\ \dot{\bar{\boldsymbol{v}}}\end{pmatrix}=\frac{1}{Z}\boldsymbol{J_t v}+\boldsymbol{J_\omega\omega} \quad (13$$

Rearranging Eq. 13 in linear form,

$$\frac{1}{Z}\boldsymbol{J_t v}=\begin{pmatrix}\dot{\bar{\boldsymbol{u}}}\\ \dot{\bar{\boldsymbol{v}}}\end{pmatrix}-\boldsymbol{J_\omega\omega} \quad (14$$

Writing Eq. 12 in compact form $A\theta=B$, observe that this Eq has a linear equation form. The focal length, the principal point, and the principal point are necessary for computing the Image Jacobian matrix but in practice, it is quite tolerant of errors in these.

3.1 Interaction Matrix for Stereo vision

Consider the pair of cameras are look at an arbitrary point in the world. The projected point in each image plane are shown by $\{p_i(x_i,y_{i)}\ ,\ i=l,r\}$ so relative equation of projecting observed points at the left and right image planes, could be written as below,

$$\boldsymbol{x_l}=\frac{X+\frac{b}{2}}{Z}\ ,\ \boldsymbol{y_l}=\frac{Y}{Z},\quad \boldsymbol{x_r}=\frac{X-\frac{b}{2}}{Z}\ ,\ \boldsymbol{y_r}=\frac{Y}{Z} \quad (15$$

Normalize the coordinates and describe Eq.15 in pixel dimensions,

$$\boldsymbol{x_l}=\frac{\boldsymbol{u_l}-\boldsymbol{u_0}}{\boldsymbol{f^*\alpha}}\ ,\boldsymbol{y_l}=\frac{\boldsymbol{v_l}-\boldsymbol{v_0}}{\boldsymbol{f^*}},\quad \boldsymbol{x_r}=\frac{\boldsymbol{u_r}-\boldsymbol{u_0}}{\boldsymbol{f^*\alpha}}\ ,\boldsymbol{y_l}=\frac{\boldsymbol{v_r}-\boldsymbol{v_0}}{\boldsymbol{f^*}} \quad (16$$

In Eq. 16 the coordinates of the camera principal point are u_0 and v_0, α is the ratio of the pixel dimensions where $\frac{dy}{dx}=\alpha$,f is the focal length and $f^*$is focal length described in pixel dimensions. Taking the time derivative of the perspective projection Equations,

$$\boldsymbol{x_l}=\frac{\dot{X}+x_l\dot{Z}}{Z}\ ,\ \boldsymbol{y_l}=\frac{\dot{Y}+y_l\dot{Z}}{Z},\quad \boldsymbol{x_r}=\frac{\dot{X}-x_r\dot{Z}}{Z}\ ,\ \boldsymbol{y_l}=\frac{\dot{Y}-y_r\dot{Z}}{Z} \quad (17$$

in the left or right image plane, the velocity of the feature point p_l could write related to the camera frame P^c as,

$$\dot{\boldsymbol{p}}=\boldsymbol{J_c^l}\dot{\boldsymbol{P}}^{\boldsymbol{c}} \quad (18$$

whereof $p_l=[p_l,p_r]$ and,

$$\boldsymbol{J_C^l}=\begin{bmatrix}\frac{\partial x_l}{\partial X} & \frac{\partial y_l}{\partial X} & \frac{\partial x_r}{\partial X} & \frac{\partial y_r}{\partial X}\\ \frac{\partial x_l}{\partial Y} & \frac{\partial y_l}{\partial Y} & \frac{\partial x_r}{\partial Y} & \frac{\partial y_r}{\partial Y}\\ \frac{\partial x_l}{\partial Z} & \frac{\partial y_l}{\partial Z} & \frac{\partial x_r}{\partial Z} & \frac{\partial y_r}{\partial Z}\end{bmatrix}^T=\begin{bmatrix}\frac{1}{Z} & 0 & \frac{1}{Z} & \frac{1}{Z}\\ 0 & \frac{1}{Z} & 0 & \frac{1}{Z}\\ -\frac{X+\frac{b}{2}}{Z^2} & -\frac{Y}{Z^2} & -\frac{X-\frac{b}{2}}{Z^2} & -\frac{Y}{Z^2}\end{bmatrix}^T \quad (19$$

the velocity of P^c related to spatial camera velocity can be write as

$$\dot{\boldsymbol{P}}^{\boldsymbol{c}}=-\boldsymbol{\omega_c}\times\boldsymbol{P^c}-\boldsymbol{v_c} \quad (20$$

solve Eq. 20,

$$\dot{\boldsymbol{P}}^{\boldsymbol{c}} = \begin{bmatrix} \dot{X} \\ \dot{Y} \\ \dot{Z} \end{bmatrix} = \begin{bmatrix} -\omega_y Z + \omega_x Y - v_x \\ -\omega_z X + \omega_x Z - v_x \\ -\omega_x Y + \omega_y X - v_Z \end{bmatrix} = \boldsymbol{\Lambda u_c} \quad (21$$

Substituting Eq. 18 in Eq.21 could write,

$$\dot{\boldsymbol{p}}_{\boldsymbol{I}} = \boldsymbol{J_c^I \dot{P}_c} \;\; \mapsto \;\; \dot{\boldsymbol{p}}_{\boldsymbol{I}} = \boldsymbol{J_c^I \Lambda u_c = J_{st} u_c} \quad (22$$

where J_{st} is the stereo-vision image Jacobian which describe the relation between a moving velocity of a camera u_c and feature point in an image $\dot{p}_I$, Considering $X = \frac{b(x_l+x_r)}{2(x_l-x_r)}, Y = y_l \frac{b}{(x_l-x_r)}$ and $Z = \frac{b}{(x_l-x_r)}$ the stereo-vision image Jacobian matrix could be written as,

$$\boldsymbol{J_{st}} = \begin{bmatrix} -\frac{a}{b} & 0 & x_l\frac{a}{b} & x_l y & -\left(1+\frac{x_l(x_l+x_r)}{2}\right) & y \\ 0 & -\frac{a}{b} & y\frac{a}{b} & 1+y^2 & -y\frac{(x_l+x_r)}{2} & -\frac{(x_l+x_r)}{2} \\ -\frac{a}{b} & 0 & x_r\frac{a}{b} & x_r y & -\left(1+\frac{x_l(x_l+x_r)}{2}\right) & y \\ 0 & -\frac{a}{b} & y\frac{a}{b} & 1+y^2 & -y\frac{(x_l+x_r)}{2} & -\frac{(x_l+x_r)}{2} \end{bmatrix} \quad (23$$

In Eq.23 $a = x_l + x_r$ which is called feature point disparity and $y = y_l = y_r$. Eventually, the stereo-vision image interaction matrix could be obtained with the velocities that expressed in the camera frame and then transformed into the sensor frame [28].

$$\dot{\boldsymbol{p}}_{\boldsymbol{I}} = \begin{bmatrix} \boldsymbol{J_l M_c^l} \\ \boldsymbol{J_r M_c^r} \end{bmatrix} \boldsymbol{u_c} = \boldsymbol{J_{st} u_c} \quad (24$$

Assume a camera spatial velocity be unit magnitude $v^T v = 1$, Due to the Eq. 24 write the camera velocity in terms of the pseudo-inverse $v = J^+ \dot{p}$ where $J^+ \in R^{2n\times 6}$ the Jacobian stack and $\dot{p} \in R^{2n}$ is the point velocities. The equation of an ellipsoid is the results of substitution these Eq's in the point velocity space.

$$\boldsymbol{p^{\cdot T} J^{+T} J^+ \dot{p} = 1} \;\; \mapsto \;\; \boldsymbol{p^{\cdot T} (JJ^T)^{-1} \dot{p} = 1} \quad (25$$

The principal axes of the ellipsoid are defined as the eigenvectors of JJ^T and the singular values of Jacobian are the radii. The condition number of JJ^T drive as the maximum to minimum radius ratio and implement the anisotropy of the feature motion. The maximum to minimum radius ratio is given by the condition number of JJ^T. this ratio indicates the anisotropy of the feature motion. The high value of condition number indicates that some of the points have low velocity in response to certain camera motions. Next section explained how to design a proper controller via selecting features.

3.2 Design Controller

The main goal of control based on vision is achieved to minimize an error that is $e(t) = p - p^*$, in this relation p is a vector of current and p^* is vectors of desired features in image space. The relation between feature variation in the time and the camera velocity are used to design a controller.

$$\dot{\boldsymbol{p}} = \boldsymbol{J_p u_c}(\boldsymbol{v_c}, \boldsymbol{\omega_c}) \quad (26$$

The robot controller uses the feature speed as input in pixel space to modify the path. Notice that the nature of p includes the visual and geometric features of the object [29]. In the first step, assume that the goal pose is fixed. So the changes in error is depend only on camera motion which is mounted on top of the end arm of 4 DOF Scara robot and the motion of camera at each time depend on robot motion. After choosing a conventional feature set for p, design a control scheme which is a velocity controller. $u_c = (v_c, \omega_c)$ is the spatial velocity of the camera, v_c is the instantaneous linear velocity of the origin of the camera frame, ω_c is the instantaneous angular velocity of the camera frame, $J_p \in k \times 6$ is called the Jacobian Matrix or feature interaction matrix. Frist assume p^* is constant,

$$\frac{\boldsymbol{dp^*}}{\boldsymbol{dt}} = \boldsymbol{0} \;\; \mapsto \;\; \frac{\boldsymbol{de}}{\boldsymbol{dt}} = \frac{\boldsymbol{d}}{\boldsymbol{dt}}(\boldsymbol{p - p^*}) = \frac{\boldsymbol{dp}}{\boldsymbol{dt}} \quad (27$$

And error just changed related to p,

$$\dot{\boldsymbol{e}} = \boldsymbol{J_e u_c} \;\; \mapsto \boldsymbol{J_e = J_p} \quad (28$$

Let assume $\dot{e} = \eta e$ so the control signal can be designed as $u_c = -\eta J_e^+ e$ where J_e^+ is the Moore-Penrose pseudo-inverse of $J_e^+ = (J_e^T J_e)^{-1} J_e^T$ where if $k = 6$ and $\det(J) = 0$, so can obtain $u_c = -\eta J_e^- e$. The desired values p^* are not constant if the observed target is moved. so the time variation of the error can now be obtained by [30].

$$= \left(\frac{\boldsymbol{\partial e}}{\boldsymbol{\partial r}}\right)\boldsymbol{u_c} + \frac{\boldsymbol{\partial e}}{\boldsymbol{\partial t}} \;\; \mapsto \;\; \dot{\boldsymbol{e}} = \boldsymbol{J_e u_c} + \frac{\boldsymbol{\partial e}}{\boldsymbol{\partial t}} \quad (29$$

The time variation of error would be created by two part, the term $\frac{\partial e}{\partial t}$ caused by the target motion which is considered as constant velocity. Controller used this term to compensate the target motion. Because in some cases the error value is estimated, therefore, the equation is rewritten as following,

$$\boldsymbol{u_c} = \hat{\boldsymbol{J}}_{\boldsymbol{e}}^{+}(-\boldsymbol{\eta e} + \frac{\boldsymbol{\partial \hat{e}}}{\boldsymbol{\partial t}}) \quad (30$$

Define a classical discrete time integral term with gain μ_I to reduce the tracking errors caused by target motion,

$$((\boldsymbol{\partial \hat{e}})/\boldsymbol{\partial t})_{\boldsymbol{k}} = \boldsymbol{\mu_I} \sum_{\boldsymbol{i=0}}^{\boldsymbol{k-1}} \boldsymbol{e_i} \quad (31$$

$(\partial \hat{e})/\partial t$ could be directly estimated based on feed-forward control, by comparing the measurements image values in each step vs. the previous one and camera velocity.

$$(\boldsymbol{\partial \hat{e}})/\boldsymbol{\partial t_k} = \frac{\boldsymbol{e_k - e_{k-1}}}{\boldsymbol{\Delta t}} - \left(\hat{\boldsymbol{J}}_{\boldsymbol{e}}^{+}\right)_{\boldsymbol{k}} (\boldsymbol{u_c})_{\boldsymbol{k-1}} \quad (32$$

In Eq.32 Δt is the control loop duration and k is the current time step. The kinematic model of the transformation

between the image features' velocities and the joints' velocities could be defined in the form of $\dot{p} = J_T \dot{q}$, where J_T is the total Jacobian, by computing the image interaction matrix featuresJ_p,velocity transformation between the camera and world coordinate frames could be shown as M_w^c and the Scara robot Jacobian in the form of J_R^w. So the total Jacobian J_T could be defined as,

$$\boldsymbol{J_T = J_p M_w^c J_R^w} \quad (33$$

Now, adjust the convergence speed of visual servoing by designing a proportional control law as,

$$\boldsymbol{\dot{q} = -\lambda \hat{J}_T^+ e} \quad (34$$

in which λ is the positive gain.

3.3 Estimate the path using the Extended Kalman filter

Visual system image processing and controller strategy calculations are time-consuming and reduce robot performance. This time delay in calculating the position of a moving object is one of the main problems in griping an object by a vision-based robot. Therefore, predicting the position of a moving object can help the system prevent such problems. Several path prediction methods have been proposed by researchers, one of which is the EKF (i.e. Extended Kalman Filter). Kalman filter is generally used to estimate the time-discrete state of linear motion of an object. With the development of the Kalman filter, the EKF has been developed and can be used for nonlinear dynamic systems. The EKF provides an approximation of the optimal estimate. The dynamics of nonlinear systems become linear in the range of the last estimated state. The Kalman sequence of the developed filter includes the following steps,

- Estimate the last state vector $\hat{X}_{k-1}$,
- Dynamic linearization of the system $X_k = f(X_{k-1}) + W_{k-1}$,around $\hat{X}_{k-1}$,
- Using Kalman filter estimation for linearized dynamic system,
- Linearization of the dynamic equation of measure $Y_k = h(X_k) + V_k$, around $\hat{X}_k$,
- Use the Kalman filter update cycle for linear measurements,

By selecting the parameters F_k and H_k as the Jacobian matrices $f(x)$ and $h(x)$ the nonlinear state dynamics for a moving object can be expressed by a linear system as follows:

$$= \begin{bmatrix} 1 & 0 & -u_k \Delta t sin(\theta_k) & \Delta t cos(\theta_k) & -\frac{1}{2} u_k \Delta t^2 sin(\theta_k) \\ 0 & 1 & u_k \Delta t cos(\theta_k) & \Delta t sin(\theta_k) & \frac{1}{2} u_k \Delta t^2 cos(\theta_k) \\ 0 & 0 & 1 & 0 & \Delta t \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (35$$
$$= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

The dynamic state equation of a moving object can be rewritten as follows:

$$\begin{bmatrix} x_k \\ y_k \\ \theta_k \\ u_k \\ w_k \end{bmatrix} = \begin{bmatrix} 1 & 0 & -u_k \Delta t sin(\theta_k) & \Delta t cos(\theta_k) & -\frac{1}{2} u_k \Delta t^2 \sin(\theta_k) \\ 0 & 1 & u_k \Delta t cos(\theta_k) & \Delta t sin(\theta_k) & \frac{1}{2} u_k \Delta t^2 \cos(\theta_k) \\ 0 & 0 & 1 & 0 & \Delta t \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{k-1} \\ y_{k-1} \\ \theta_{k-1} \\ u_{k-1} \\ w_{k-1} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \zeta_u \\ \zeta_w \end{bmatrix} \quad (36$$

$$\begin{bmatrix} x_k \\ y_k \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_k \\ y_k \\ \theta_k \\ u_k \\ w_k \end{bmatrix} + \begin{bmatrix} \gamma_x \\ \gamma_y \end{bmatrix}$$

The equations of the EKF algorithm are expressed as follows,

- Estimation or prediction phase

$$\hat{X}_k^* = f_{k-1}(\hat{X}_{k-1}) \approx F_{k-1} \hat{X}_{k-1}, P'_k \quad (37$$
$$= F_{k-1} P_{k-1} F_{k-1}^T + Q_{k-1}$$

- Filtering phase

$$\hat{X}_k = \hat{X}_k^* + K_k [Y_k - h_k(\hat{X}_k^*)] \implies \hat{X}_k \quad (38$$
$$= \hat{X}_k^* + K_k [Y_k - H_k]$$
$$K_k = P'_k H_k^T [H_k P'_k H_k^T + R_k]^{-1} , \quad (39$$
$$P_k = P'_k - K_k H_k P'_k$$

3. Simulation and Experimental results

Defining a properly detailed framework of the stereo visual servo is necessary for achieving precise computer simulation. In this section, the schematic model based on the eye in hand stereo visual servoing system is expressed. The base method is shown in the Fig. 1. In order to determine the coordinates of the parts relative to the robot gripper, the cameras on the robot take images of the robot's workspace online. Therefore, to estimate the speed and location of the piece in pixels world, the path of the robot arm can be plotted with the least square error. In order to automate the assembly of product lines, a special assembly machine was designed and built.

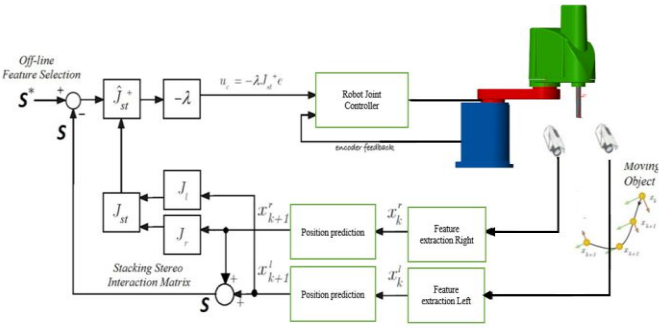

Figure 1: prediction of object positions in stereo image-plane to grasp a moving target by Scara robot

This special machine includes a Scara robot with 4 degrees of freedom and an assembly table for placing parts. In order to optimize and reduce the error, two cameras are installed on the end arm of the robot, which acts as control feedback. To implement the theory presented in the paper, it was decided to first simulate the single-camera mode based on the image. As you can see in Figure 2, the camera mounted on robot sees the four corners of a square and directs the gripper controller to reach predefined points in the image space.

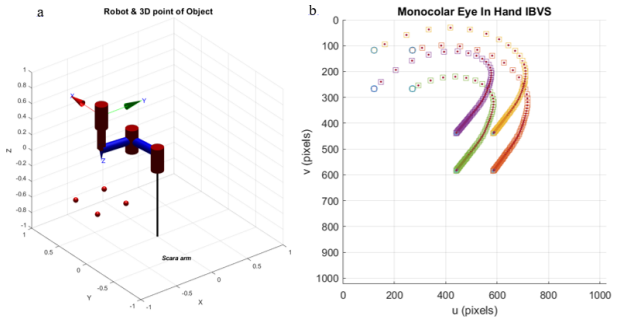

Figure 2 monocular IBVS task a- Image point trajectories b- Experimental robot simulation 3D-model and 3D point

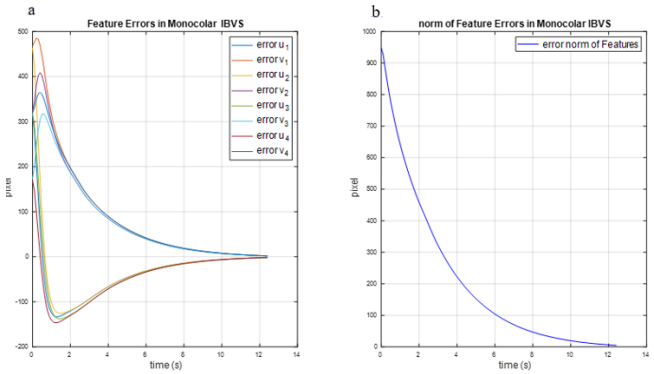

Figure 3: Monocular IBVS task a- Feature error b- The norm of feature error

As shown in Figure 3, it takes at least 12 seconds for the gripper to reach the body using the visual control information. The reason for the time-consuming process is that in the case of a single camera, it is not possible to estimate the distance of the object to the gripper at any time, and therefore the image interaction matrix is not updated.

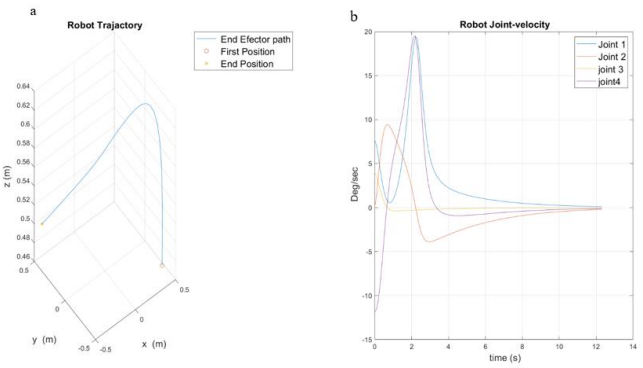

Figure 4: a –Griper of robot trajectory in 3-D space, b- Robot joint velocity components in a monocular IBVS task

Stereo interaction matrix concludes from the transformation of a matrix of the left and right camera frame to the virtual sensor frame and then to control of robot we have to transform sensor frame to the end-effector frame. The cameras were calibrated using the MATLAB apps and drive the intrinsic and extrinsic camera's parameters. Rectification is used in stereo vision to simplify the problem of finding matching points between images. The 3-D location of points viewed by Cam Left and Cam Right could be obtained having the stereo Calibration parameters such as intrinsic matrix, essential matrix. The Essential matrix relates the image of a point in one camera to its image in another camera, given a translation and rotation$P_0^T E P_1$. This matrix (E), could be calculated by knowing the pose between the two views, or a set of known point's correspondences. The matrix E could be obtained, as, $E = [t_x]R$ in which $[t_x]$ is the skew-symmetric matrix associated with the vector t and (R, t) is the rigid body transformation between two cameras. Due to the time-consuming control process based on stereo vision, we used estimation prediction methods to improve the speed of the control process. Three estimation methods include a Kalman filter, Extended Kalman filter and Recursive Least Square method were used as prediction methods and compared. For investigation of the performance of the approach, the motion of the object was selected manually and randomly. Due to the mechanical limitation of the robot we used limited relative speed to the movement of the object. As mentioned, the robot moves and aligns its end-effector with respect to the object position to perform grasping task by pre determinant image features.

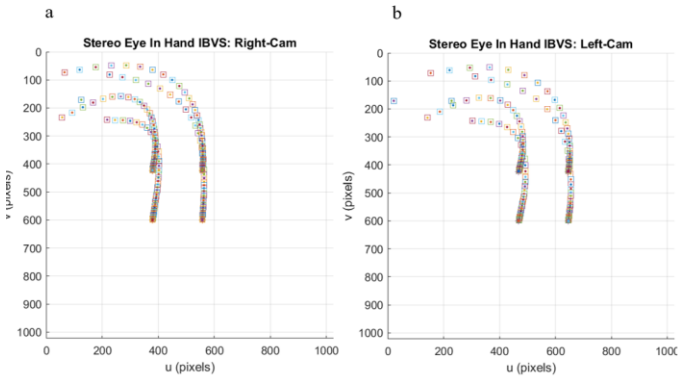

Figure 5 Stereo IBVS image feature trajectories for (a) Right image plane and (b) Left image plane.

By comparing the feature trajectory produced in the image plane, the stereo with monocular methods, it is clear that the paths produced by the stereo approach are completely smoother and more efficient. Feature errors in left and right images are illustrated in Figure 6 (a, b) the rate of decreasing error is higher and oscillations and overshoot of it was lower compared to the monocular case. Analyzing the simulation results, it is quite clear that the stereo mode performed better than the single-camera, but again the time of the gripper to reach the target is still high.

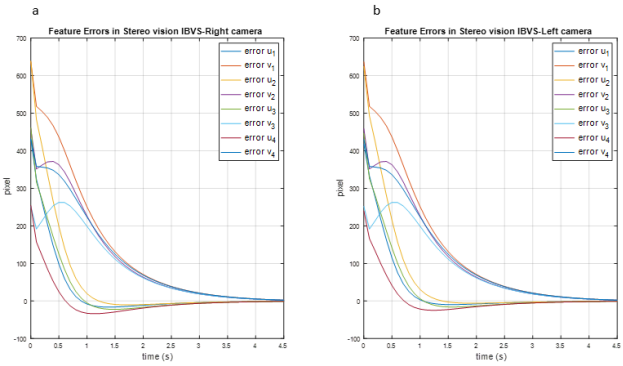

Figure 6 The Stereo IBVS task feature errors in (a) left and (b) right images

As shown in Fig 7-a norm of feature errors in left and right images in the Stereo IBVS task is converge to zero in 4.5 seconds. In this case, Camera velocity components is shown in

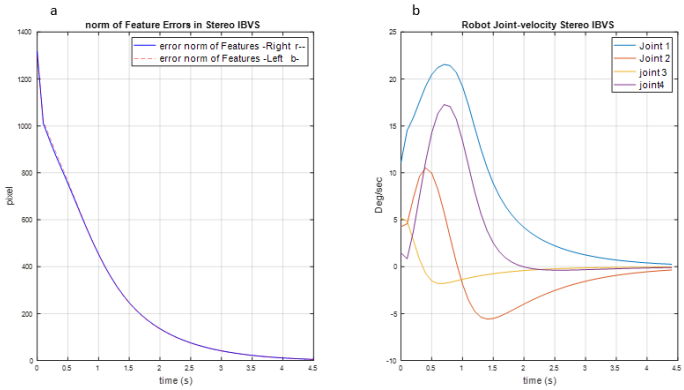

Figure 7 -b. As mentioned in the literature review, comprising the results shown that in the case of the monocular vision, the overshoot of system is large while convergence speed is slow. Therefore, by adding prediction algorithms, robot performance can be improved and time can be reduced. Therefore, by using the current position of the points as well as the speed of the object in the image space, the position of the target movement path is predicted, and by optimizing the gripper to reach the goal, the process of getting the workpiece can be reduced. The effect of different prediction methods such as Kalman filter, RLS, and EKF on the performance of stereo-based visual servoing has been investigated and the advantages and disadvantages of each method have been stated.

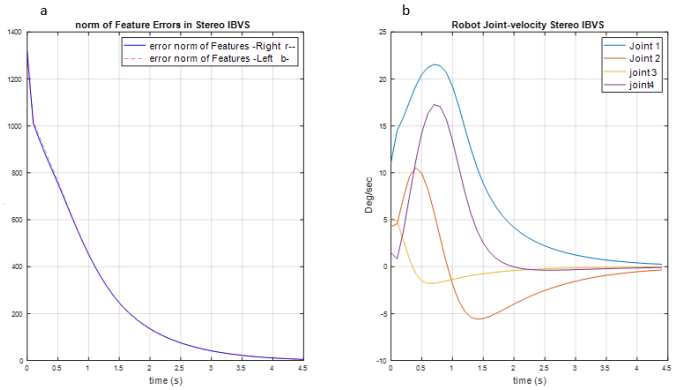

Figure 7 norm feature errors in left and right images. b-Camera velocity components in a stereo eye in hand IBVS task

The path of the object in the image space is often non-linear, so the Kalman filter in many cases does not meet the needs of the problem except under certain conditions. The effect of the least recursive squares on S-IBVS is shown in Figure 8.

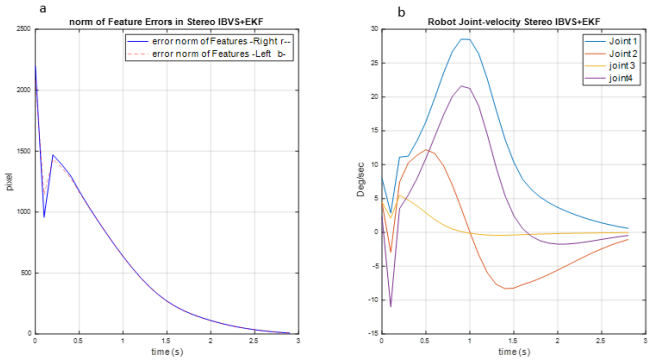

Figure 8 .a- Feature error norm of Stereo IBVS+EKF b- Camera frame velocity components Robot end-effector Stereo IBVS+EKF task

As you can see in Figure 8, the process time is about 15% optimized compared to the unused state, since this algorithm moves the robot end effector more quickly towards the target when the difference between the points observed in the image space and the desired point is large. By adding the extended Kalman filter, the simulation can be implemented for all linear and nonlinear points and modes. According to the simulation results in Figure 9, the extended Kalman filter method improves robot performance and increases operating time by at least 20%. The stereo visual servoing has fast convergence and small overshoot compare to monocular visual servoing, because of updating the depth information at every step by image interaction matrix. As a result of this, the robot is more stable due to the correct feedback comment.

Table 1. The brief results of tracking and grasping tasks in IBVS cases

IBVS - CASE	Convergence time (s)	angular velocity at t=0 (deg/s)	linear velocity at t=0 (deg/s)	tracking error (pixel)
Monocular IBVS	13.8	18	95	170
stereo IBVS	4.5	25	70	115
stereo IBVS +Kalman	3.6	8.5	46	55
stereo IBVS +EKF	2.8	5	38	35
stereo IBVS + RLS	3.8	10	69	73

As a result of this, the robot is more stable due to the correct feedback comment. In stereo approach, the object motion was modeled in both image planes. By using it the trajectory and position of the robot end-effector could be predictable. According to Equation (33) in an image-based control scheme, the stereo Jacobian matrices are used to re-position the image features to the desired positions. The effect of using the extended Kalman filter and the Recursive least squares are compared in Figure 8.

The summarize of results for tracking and grasping cases gathered in table 1. the parallelism of the camera in the stereo IBVS cause smaller sensor frame velocities at beginning that indicate more efficiency of parallel system compare to non-parallel. To evaluate the robot's motion stability and target traceability using image features, the robot gripper was placed in different positions relative to the target object. By examining the different positions of the target relative to the end effector at near and away distances, the results showed that the use of path optimization algorithms in all situations reduces the robot movement time and optimizes the robot movement path. Logitech Pro C920 cameras were used, and a fixed focal length was taken to obtain the depth or distance of the object from the gripper. The stereo camera online grabs a pair of images of the workplace. To Reduce the error in detection, images are filtered by 2D discrete Wavelet [31] to eliminate noise from the lighting as well as camera movement. Chose a moving object as shown in Fig. 9. First of all, filter the images and sharpening the edges and corners. Match the object in left and right images. Detect the object and extract features. The scene is captured by the cameras at every moment, and according to the estimation method, the location of the moving target is estimated at the next moment. This information is given to the robot controller to get the workpiece by the robot.

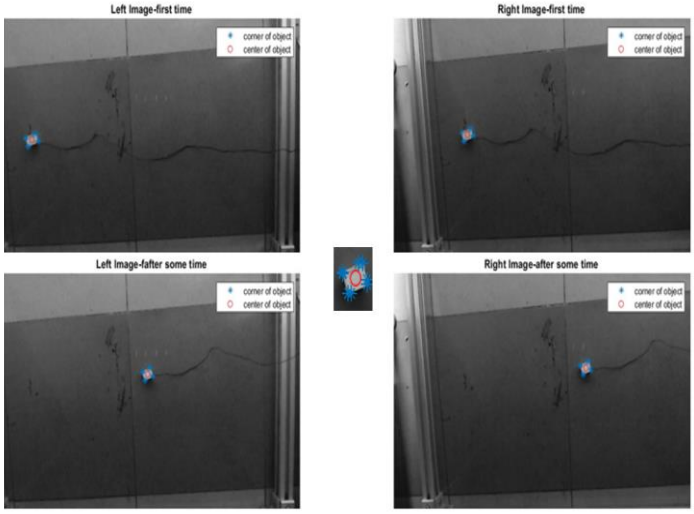

Figure 9: Detecting object and matching the right and left image for obtaining 3-D location of Objects a- at first time b- after some time.

Another scene was selected experimentally as shown in Fig. 10 where the center of the piece and the surrounding holes were selected as the constraint. Therefore, by using these constraints, it can be obtained the parts orientation and the amount of rotation required by the robot end effector to be assembled can be achieved with good accuracy.

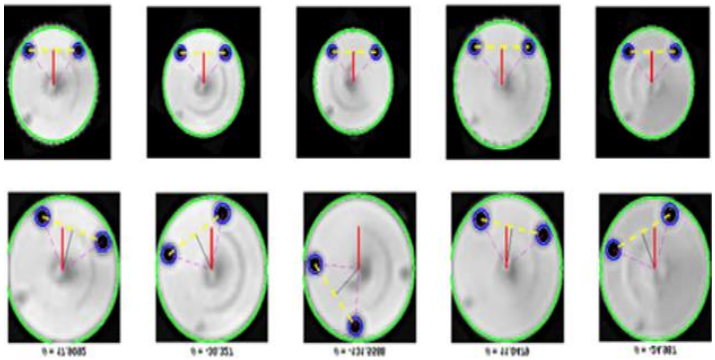

Figure 10: Finding orientation of parts for accurate assembling using Stereo Image-based visual servoing.

According to the algorithms used in this paper, the distance and location of the parts compared to the gripper are calculated according to the triangulation, and the errors of the position and the orientation of parts are minimized by applying the estimation methods.

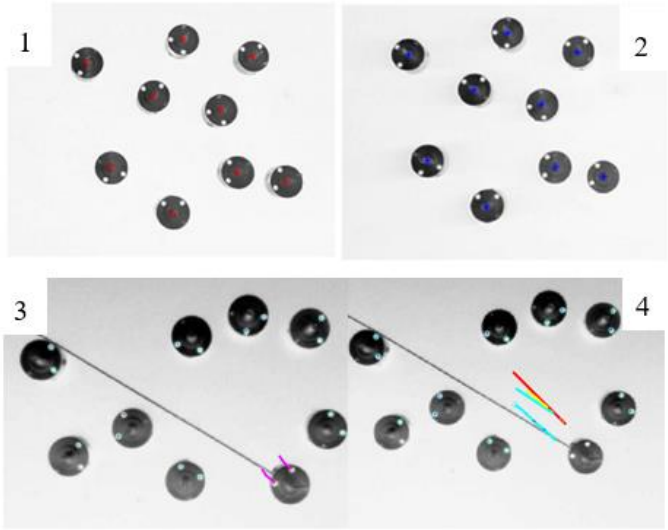

Figure 11: Prediction of component position and moving direct

Figure 12: a. Malek Ashtar University (4-dof) Scara robot designed and manufacture for tiny- fine parts assembling b-. Scara robot with Stereo_ IBVS (C922 Logitech PRO camera) mounted on the end effector

In order to confirm the proposed method, it was decided that after testing the successful simulation, the design and construction of the Scara robot with the aim of assembling delicate and precise parts were put on the agenda. Based on the preview method, the cameras were placed on the robot's end-effector, and the robot's controller uses the feedback of the resulting image information to carefully assemble the parts. The capability of proposed visual servoing control system and its tracking algorithms in real varies task was tested and verified. For experimental tests, a Scara robot shown in Fig. 13-a was used which is made to assemble delicate and precise parts. As shown in Fig. 13-b two Logitech C922 PRO Cam's are attached on the end effector. These cameras detect the distance and location of parts and provide information to the robot controller for processing and decision making. The simulation and experimental results show that image feature trajectories in the systems with Kalman filter and EKF are almost the same. The behavior of them are quite smoother than the system with RLS estimator. Using the Extended Kalman filter as an estimator give better convergence performance compare other previously mentioned methods. In addition, the end-effector trajectory in 3-D space has less unnecessary motion from the starting point to the catching position in the case that predict via EKF method.

4. Conclusions

In order to automate the assembly of product lines, a special assembly machine was designed and built. This special machine includes a Scara robot with 4 degrees of freedom and an assembly table for placing parts. In order to optimize and reduce the error, two cameras are

mounted on the end arm of the robot, which acts as control feedback. The results show that stereo-IBVS had more smooth and fast response compare to the monocular system. The trajectory of point's path planning in image planes at stereo smooth and the camera velocity don't include large oscillation. Adding prediction techniques to the proposed algorithm, helps the robot to track objects quickly with fewer convergence errors. The proposed algorithm makes it possible to identify the main features of the object so that by having this information online, the robot can detect the position and orientation of the part by feedback information from the image processing system. To test the proposed method, several prediction algorithms were added to the stereo vision controller. The results showed that when the Extended Kalman filter estimator algorithm is used, time of the target tracking and grasping compared to other methods is about 20 percent less. It is inferred from the results that, it is inferred from the results that the case with the EKF estimator has better behavior in end-effector 3-D trajectories and shows better tracking and convergence performance. The camera velocity components in the system with EKF in comparison with the system with the RLS estimator start with relatively lower values. The method was implemented on the Scara robot, and the robot was able to perform the task using the feedback information of the image processing system. For future work, it is suggested to use new prediction methods such as reinforcement neural network or other control methods like fuzzy-PID controller or nonlinear control methods such as sliding mode control. For designing controller acceleration control can be considered instead of speed control.

Reference

[1] M. Abdelhamid, J. Beers, and M. Omar, "Extracting Depth Information Using a Correlation Matching Algorithm," 2012.

[2] R. He, J. Rojas, and Y. Guan, "A 3D Object Detection and Pose Estimation Pipeline Using RGB-D Images," 2017, [Online]. Available: http://arxiv.org/abs/1703.03940.

[3] A. Yamashita, T. Kaneko, S. Matsushita, K. T. Miura, and S. Isogai, "Camera Calibration and 3-D Measurement with an Active Stereo Vision System for Handling Moving Objects," *J. Robot. Mechatronics*, vol. 15, no. 3, pp. 304–313, 2003.

[4] P. Corke, *Robotics, Vision and Control: Fundamental Algorithms In MATLAB ®Second, Completely Revised*, vol. 118. Springer, 2017.

[5] A. Mohebbi, "Real-Time Stereo Visual Servoing Of a 6-DOF Robot for Tracking and Grasping Moving Objects," 2013.

[6] H. Wang, "Adaptive visual tracking for robotic systems without image-space velocity measurement," *Automatica*, vol. 55, pp. 294–301, 2015.

[7] K. Rahardja and A. Kosaka, "Vision-based bin-picking: recognition and localization of multiple complex objects using simple visual cues," in *IEEE International Conference on Intelligent Robots and Systems*, 1996, vol. 3, pp. 1448–1457, doi: 10.1109/iros.1996.569005.

[8] A. Astolfi, L. Hsu, M. S. Netto, and R. Ortega, "Two solutions to the adaptive visual servoing problem," *IEEE Trans. Robot. Autom.*, vol. 18, no. 3, pp. 387–392, 2002.

[9] Y. Shirai and H. Inoue, "Guiding a robot by visual feedback in assembling tasks," *Pattern Recognit.*, vol. 5, no. 2, pp. 99–108, 1973.

[10] F. Chaumette, "Potential problems of stability and convergence in image-based and position-based visual servoing," in *The confluence of vision and control*, Springer, 1998, pp. 66–78.

[11] D. B. Westmore and W. J. Wilson, "Direct dynamic control of a robot using an end-point mounted camera and Kalman filter position estimation," in *Proceedings. 1991 IEEE International Conference on Robotics and Automation*, 1991, pp. 2376–2377.

[12] A. Ghasemi, P. Li, and W.-F. Xie, "Adaptive Switch Image-based Visual Servoing for Industrial Robots," *Int. J. Control. Autom. Syst.*, pp. 1–11, 2019.

[13] W. Feng *et al.*, "Inertial measurement unit aided extrinsic parameters calibration for stereo vision systems," *Opt. Lasers Eng.*, vol. 134, p. 106252, 2020.

[14] W. Li, C. Song, and Z. Li, "An accelerated recurrent neural network for visual servo control of a robotic flexible endoscope with joint limit constraint," *IEEE Trans. Ind. Electron.*, 2019.

[15] K. Hashimoto, T. Ebine, and H. Kimura, "Visual servoing with hand-eye manipulator-optimal control approach," *IEEE Trans. Robot. Autom.*, vol. 12, no. 5, pp. 766–774, 1996.

[16] F. Chaumette and E. Malis, "2 1/2 D visual servoing: a possible solution to improve image-based and position-based visual servoings," in *Proceedings*

2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No. 00CH37065), 2000, vol. 1, pp. 630–635.

[17] Z. Jia *et al.*, "Improved camera calibration method based on perpendicularity compensation for binocular stereo vision measurement system," *Opt. Express*, vol. 23, no. 12, pp. 15205–15223, 2015.

[18] H. C. Radhakrishnamurthy, P. Murugesapandian, N. Ramachandran, and S. Yaacob, "Stereo vision system for a bin picking adept robot," *Malaysian J. Comput. Sci.*, vol. 20, no. 1, pp. 91–98, 2017.

[19] R. Tsai, "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses," *IEEE J. Robot. Autom.*, vol. 3, no. 4, pp. 323–344, 1987.

[20] J. Heikkila, "Geometric camera calibration using circular control points," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 10, pp. 1066–1077, 2000.

[21] S. Gai, F. Da, and X. Dai, "A novel dual-camera calibration method for 3D optical measurement," *Opt. Lasers Eng.*, vol. 104, pp. 126–134, 2018.

[22] P. Monasse, J.-M. Morel, and Z. Tang, "Three-step image rectification," in *BMVC 2010-British Machine Vision Conference*, 2010, pp. 81–89.

[23] W. Chen, Y. Lu, B. Yang, G. Ma, Z. Wang, and Y.-H. Liu, "Automatic field of view control of laparoscopes with soft RCM constraints," in *2018 13th World Congress on Intelligent Control and Automation (WCICA)*, 2018, pp. 653–658.

[24] F. Janabi-Sharifi and M. Marey, "A kalman-filter-based method for pose estimation in visual servoing," *IEEE Trans. Robot.*, vol. 26, no. 5, pp. 939–947, 2010.

[25] J. Park, J. Yun, and J. Lee, "Trajectory estimation of a moving object using Kalman filter and Kohonen networks," *Robotica*, vol. 25, no. 5, pp. 567–574, 2007.

[26] P. K. Allen, A. Timcenko, B. Yoshimi, and P. Michelman, "Trajectory filtering and prediction for automated tracking and grasping of a moving object," in *Proceedings 1992 IEEE International Conference on Robotics and Automation*, 1992, pp. 1850–1856.

[27] R. E. Kalman, "A new approach to linear filtering and prediction problems," *J. basic Eng.*, vol. 82, no. 1, pp. 35–45, 1960.

[28] P. Martinet and E. Cervera, "Stacking jacobians properly in stereo visual servoing," in *Proceedings 2001 ICRA. IEEE International Conference on Robotics and Automation (Cat. No. 01CH37164)*, 2001, vol. 1, pp. 717–722.

[29] F. Chaumette and S. Hutchinson, "Visual servo control. I. Basic approaches," *IEEE Robot. Autom. Mag.*, vol. 13, no. 4, pp. 82–90, 2006.

[30] D.-J. Kim, R. Lovelett, and A. Behal, "Eye-in-hand stereo visual servoing of an assistive robot arm in unstructured environments," in *2009 IEEE International Conference on Robotics and Automation*, 2009, pp. 2326–2331.

[31] M. Jeddi, A. R. Khoogar, and A. Mehdipoor Omrani, "Reducing Image Size and Noise Removal in Fast Object Detection using Wavelet Transform Neural Network," *ADMT J.*, vol. 13, no. 2, pp. 13–21, 2020.

Biography

Mahmoud Jeddi is currently Graduated with a PhD in Mechanics from the department of Mechanical Engineering at the Malek Ashtar University of Technology, Tehran, His current research interests include Robotic, Control and Vision.

Ahmad Reza Khoogar is Associate Professor of Mechanical Engineering at the Malek Ashtar University of Technology, Tehran, Iran. He received his PhD in Mechanical engineering from The University of Alabama, USA in 1989. His current research interests include Robotic, Control and Artificial Intelligence.

Ali Mehdipoor Omrani is Associate Professor of Mechanical Engineering at the Malek Ashtar University of Technology, Tehran, Iran. He received his PhD in Mechanical Engineering from K. N. Toosi University of Technology at 2007. His research interests include control of distributed parameter systems and robotics.