Eye in hand robot arm based automated object grasping system

Asnor Juraiza Ishak¹, Sarmad Nozad Mahmood²
¹, ²Department of Electrical and Electronic Engineering, Faculty of Engineering, Universiti Putra Malaysia

ABSTRACT

The modern robotic systems state that the tracking methodology and the visual servoing are imperative to discover the existence of an object and excite the robot in order to manipulate the target. This paper shows a new object tracking and grasping technique in real time based on Eye in Hand visual servoing structure via a camera mounted at the end of the robot arm. The working principle of the robotic system depends mainly on the prediction based on Kalman filter method that estimates the next location of a moving object in order to specify the path of the target under the scope of the camera. Hereby, the proposed system observes the object and studies its behavior based on the previous state in order to grasp the target at the exact position. Furthermore, the vision system implements feedback control approach to keep the extracted information of the object updated to solve the stability and the reliability issues that might be encountered. It has to be mentioned that the proposed robotic system was tested by grasping moving objects in different speeds and directions. In addition, the grasping of a stationary object was tested to confirm the practical and the theoretical results. As a final result, it can be stated that the speed of the object is directly proportional with the grasping time and vice versa.

Keyword:
Eye in hand robot arm
Eye to hand robot arm
Visual servoing system
Object tracking
Kalman filter method

Corresponding Author:
Sarmad Nozad Mahmood
Department of Electrical and Electronic Engineering, Faculty of Engineering
Universiti Putra Malaysia
Serdang, 43400 Seri Kembangan, Selangor, Malaysia
Email: sarmadnmahmood@gmail.com

1. Introduction

The industrial activities based on robotic systems and robot arms were implemented considerably in the recent decade. The tendencies towards system automations and reconfigurable manufacturing systems have excited control modules to invent and introduce unique techniques with marvel approaches. The idea of integrating the robotic system with computer vision has proposed in wide areas recently. There exist a lot of civil, military, and industrial applications depend drastically on the implementation of such invention to activate system automations in different fields. The approach allows the robotic system to take decisions based on the image captured for the situation by the vision system (camera). The methodology of the work starts by equipping a camera on the end of a robot arm to create what is called Eye in Hand (EIH) [1, 2, 3]. The most suitable fields to employ and develop such systems are identified by the industrial, medical, military and reconfigurable manufacturing systems. In the medical area, the surgical operations can be performed remotely
such that the surgeon instructs the robot arm for moving to different locations during the surgery and apply actions in specific places on the human body depending on the image captured by the camera. Furthermore, the anti-collision systems, the pre-collision the brains, unmanned vehicles and the obstacle avoidance can be integrated by the proposed robotic system to be implemented in the military fields for example controlling a robot vehicle remotely inside perilous environments that the human being are prohibited to reach. In addition, the proposed robotic system can be combined completely with the reconfigurable manufacturing systems that apply system automation idea to move, equip, disassemble, polish, weld, and assemble the sections of the cars based on EIH structure with respect to supervisory control and data acquisition SCADA system or discrete event system DES. All the presented fields act properly under the domination of image processing methodology due to the major dependency on camera’s vision. As an introduction, the processing procedure performed on a certain image enhances the portrait and eliminates the noise and the demerits based on several operations. Image processing principle is related with the object tracking opportunity to realize the information required to specify the exact position of an object. It is worth mentioning that the appropriated means to track objects are identified by ultrasound waves, infrared, camera, etc. The amounts of the information that stored in a memory and analyzed later are considered as a passive tracking method. Whereas, tracking objects by its traditional meaning depending on specific information is called active tracking method. The targets that are tracked can be even directly such as tracking moving objects or based on video surveillance such as monitoring status at specific location. The tracking processes are illustrated briefly by mining data related from an interested image edge in order to discover the shape of the object, tracing the object among the frames taken from the image, and evaluating the target object to understand its behavior [4, 5, 6]. The hierarchical structure that integrates the robotic systems, the edge processing and the tracking methods in a uniformed robust module is called visual servoing. This technique manages the motion of the arm by considering the information realized from the vision unit. The robots that are directed to run in a tricky environment need to be under the supervision of the visual servoing system. The vision system is a camera mounted at the end of a robot arm in order to create the eye that sees the motions in the workspace depending on image processing operations that are performed on the captured image frames. The location of the camera allows the robot to have free space better by more than 4 degrees. Furthermore, the information taken from the proposed image is transported and stored at the computer through the universal serial bus port in order to apply image processing operations on the image and to estimate the updated location of the target object using Kalman filter algorithm [7, 8, 9]. Moreover, the interfacing between the SSC-32 servo controller of lynx robot arm and the computer is done via the serial port that provides pulse width modulation PWM signals to activate the servo motors available in the arm with respect to the angles. Finally, the proposed robotic system was examined and verified to confirm the applicability of the system when it intends to grasp object moving in different speeds cm/sec.

2. Visual Servoing

The servoing procedure extracts the information of an image captured by the camera that monitors robot arm scope based on motion activities. Visual servoing methodology can be adapted depending on camera location. Servoing methodology is categorized into two aspects identified by eye in hand (EIH) and eye to hand (ETH) methods as shown in Figure 1. In eye in hand approach, the robot arm body and the camera were integrated by affixing a camera on arm’s end in order to award free synchronous motion to the entire piece. Whereas, eye to hand methodology states that the object scope is proposed to be monitored by a stationary camera affixed somewhere such that it can works as a parallel composition with respect to object motion and robot arm grip. Eye in hand methodology guarantees free mobility, high accuracy, and exact tracking positions, however the scope field would be narrow. In the other hand, eye to hand method allows wide tracking area due to the unique monitoring technique that tracks the object and its corresponding robot arm accordingly [1, 10]. The Visual servoing possesses an aspect is that defining the signal error using feedback control blocks depending on Image Based Visual Servoing and Position Based Visual Servoing denoted by IBVS and PBVS respectively. The signal error that improves the controllability of the system can be realized by IBVS with respect to the image captured and the instructions ordered to the robot. Furthermore, the PBVS specifies the position of the camera depending on the options taken from a corresponding image [11].
3. Tracking Methods

Object tracking is considered the most appropriate problem to study image processing field. Object tracking techniques are developed considerably during the last decades. On the other hand, the whole matter remained the trickiest problem that needs to be cracked. The significance of object tracking is proposed by various applications such as human computer interaction, video observation, and robot navigation. In this paper, both tracking and detection techniques were studied depending on well-known algorithms given by Kalman filter technique.

4. Approach Based Kalman Filter

NASA agency recommended a lot of authors to work and develop object tracking methods. In 1960, Kalman R.E presented a new way to track the objects and hence the method labeled by his name. Kalman technique presents a method for tracking and describing solutions based on a unique algorithm. The proposed algorithm has been assigned to be the main topic that most of researchers built their theories and solutions depending on the terms and the conditions given by this technique particularly in the fields of automatic controlling and navigations [7, 12] Kalman filter, Extended Kalman filter, Unscented Kalman filter. Kalman filter demonstrates several mathematical computations to realize the least mean squared error. It is worth mentioning that the algorithm implements the data realized from the last step in the next step by taking into consideration the unknown system models. Kalman filter plays a great roll in various applications such as site estimation, electric potential, velocity appreciations and temperature degrees. This paper concentrates on object tracking that is considered the field that Kalman algorithm depends on. Kalman filter acts well in the field of state space area in order to expect the new updated elements effectively that is the algorithm describes the dynamical state systems that includes the noise. With respect to the state space areas, Kalman filter is constructed from two stages identified by the prediction stage and refine stage as shown in the block diagram in Figure 2. The dynamic mode of the system can be estimated by the prediction stage, while the correction stage notices the whole module to shrink the possibilities of the errors [13].
Kalman filter algorithm works depending on three tuples identified by observation model, dynamic model and state vector. The variables used in the state vector are employed to describe system dynamic. The states can be represented by orientation, position, speed, etc. Kalman filter tracking algorithm predicts the location of the object that moves in reality under camera’s scope. The realized information regarding the proposed object such as the orientation, center etc., will be kept in order to guide robot arm motion accordingly. Moreover, the information based Kalman filter algorithm with respect to the captured image defines $X, Y$ axis values and the state $k$ in order to specify the current location of the object as follows [14]:

$$X_k = \begin{bmatrix} X_k \\ Y_k \end{bmatrix}$$  \hspace{1cm} (1)

Furthermore, the proposed vector is represented by two parameters identified by the predicted and refined values. In addition, the conversion processes of the proposed vector with respect to the time can be the key task of the dynamic model. While, the observation model shows the mapping conditions among the measured states, which approves the linearity of the system as follows:

$$L(t) = H . X(t) + \omega(t)$$  \hspace{1cm} (2)

The Kalman filter algorithm starts defining several parameters to satisfy the general form of the covariance matrix. These parameters were identified by predicted noise $P$, measured noise $R$, conversion item from measured state matrix $H$, and state conversion matrix $\Phi$. Firstly, the robot arm moves synchronously to satisfy the defined $X, Y$ location depending on EIH working principle. However, the defined value initially is not exists that is the reason to assess the predicted value as much as the measured value that represents the current location of the object. Next, the robot arm reaches the predicted location then new parameters calculations start in order to fulfill the prediction stage. Last but not least, the robot arm moves to the predicted location and excites the refined stage to initialize the processes of adapting the predicted values alongside with the measured values of the object [8, 15]. For more clarity, the refined and prediction stage algorithm is given as follows:

- **Prediction stage**

$$X'_k = \Phi x_{k-1} + B u_{k-1}$$  \hspace{1cm} (3)

$$Z'_k = H x_k + v_k$$  \hspace{1cm} (4)

Where, $k$ is the current iteration, $X'_k$ is the state vector in discrete time, any $k - 1$ value means the previous state, $\Phi$ is $(n \times m)$ constant matrix, $u$ represents the control vector, $B$ is the optional control matrix, $Z$ is the measured value which is usually given as vector, $H$ is the measured state matrix, and $v$ represents the noise vector. Moreover, the proper filter gain for the pre-predicted error matrix is computed as follows:

$$P'_k = \Phi P_{k-1} \Phi^T + Q$$  \hspace{1cm} (5)

Where, $P'_k$ represents the error matrix, $P_{k-1}$ is the predicted error in the previous state, and $Q$ is the noisy matrix that shows the flexibility of the signal to be altered.

- **Refined stage**

$$K_k = P'_k H_k^T [H_k P'_k H_k^T + R_k]^{-1}$$  \hspace{1cm} (6)

Furthermore, $K_k$ represents the filter gain, $H_k$ is the observation matrix, $R_k$ is the zero average noise array, which presents the amount of the errors in the module. In addition, the predicted state vector can be computed as follows:

$$\hat{X}_k = X'_k + K_k [Z_k - H_k X'_k]$$  \hspace{1cm} (7)

Finally, the covariant matrix that estimates the amount of the predicted errors can be expressed as follows:

$$P_k = P'_k - K_k H_k P'_k$$  \hspace{1cm} (8)
5. **Object Detection**

The detection scenario starts by capturing a snapshot for the object using the camera in order to obtain the digital image of the object. The digital image is constructed from numerous pixels that limit the quality of the image with respect to the color. The pixel is the smallest picture element in an image where the picture can be an Red Green Blue (RGB) or gray scale form. The representation of the two image forms is different than each other and hence the processing over the images will be performed in a different way. An image is constructed from three colors identified by (Red, Green, and Blue) and the combination among these three colors results the other spectrum colors. In order to complete object detection processes, the detected object image should be in gray scale form. The gray scale image is composed of two dimensional functions denoted as \( f(x, y) \) where \( x \) and \( y \) represent the coordination of the image. To be more specific, the RGB form can be converted to gray scale as follows:

\[
\text{Gray Scale} = 0.3 \times R + 0.59 \times G + 0.11 \times B
\]

Where, \( R \), \( G \), and \( B \) represent the colors Red, Green, and Blue respectively. The gray scale form allows a lot of techniques to get an object out from a background frame in order to simplify the image processing methodology with respect to the filtering procedures. The median filter is considered the most suitable filter that processes the central pixel of an image. Moreover, another operation can guarantee more techniques to extract image feature is called morphological operation. The morphological operation \( MO \) is a theory that presented to analyze the geometric form of the geologic sections. The operation employed in various applications such as video observation and tracking tasks. The operation is applied widely in the cases of processing by extracting the elements of an image to collect the description for the corresponding area in a shape. In addition, the morphological operation eliminates the undesired regions in a shape. It is worth mentioning that erosion process is considered as an important option allowed by this technique such that the erosion results a processed image with a small size and a demerit feature. Whereas, the dilation violates the erosion by generating a typical image with an enhanced characteristics [15, 16].

6. **Robot Arm Kinematics**

The necessity for robotic systems increased drastically in the last decade due to robot contribution in various applications such that industrial application, system automations, reconfigurable manufacturing systems, etc. In addition, the simplicity of the implementation and the setup had increased the demand to include the robotic systems in various fields in order to reach the desired production with less effort. Most of researchers studied robot bodies while the huge trend was towards the arm of the robot. Robot arm plays a great roll in the industrial applications especially in the fields that depend predominantly on the programmable systems to run their work. The discrete event systems DESs and the SCADA systems can be the suitable example on such robot interaction areas [17]. The reason to concentrate on the robot arm is that the arm is constructed from several joints, which are considered the central region of the activities. The proposed joints are controlled by using motors in order to reach free motion different than ordinary human arm due to the ability to move in 360 degree to cover the full directions with different speeds. The gripper is attached to the end joint of the arm, which can hold the target with respect to movement interaction. This section concentrates on controlling the mechanism structure of the robot arm depending on the joints and the links. It can be stated that the main governor of arm movement is the sensors and the actuators. The target is tracked depending on the actuators that activate the robot arm to adapt its motion with the speed of the object. Whereas, the sensor satisfies the availability of the object even in camera's scope or gripper filed. The increase in the number of joints may result a robot arm with more flexibility and mobility. There are two kinds of joint in a robotic arm such as revolt and prismatic joints. The robot kinematics can be concluded based on these joints with respect to degree of freedom (DOF) structure. It has to be mentioned, that the revolt joint are more applicable than the other joint due to the free rotational movement between two links that is allowed by revolt joint. In this work, the employed arm is Lynx robot arm which possess five DOF as same as the human being arm with respect to the gripper. This means, that the proposed robot arm is constructed of 5 DOF and 1 gripper as shown in Figure 3. The robot arm controller is provided by the signal from the PC via the corresponding cable used for compatibility issues.
There are several motions in each part of the robot such as the shoulder is considered as the base that controls the motion of the arm [18]. Furthermore, the elbow is another base that dominates the motion of the wrist up and down or rotates the wrist to the full directions. For more simplicity, the presented joints can be illustrated clearly by specifying a symbolic model for each joint $J$ with respect to their rotational angles as shown in Figure 4.

The joint $J_1$ represents the base which moves with respect to the axis $z_1$. Moreover, $J_1$ generates the angle of the rotation $\theta_1$ that rotates around $z_0$ in $x_0 y_0$ plane. The joint $J_2$ represents the shoulder that moves alongside with the vertical axis of $J_1$, knowing that the corresponding rotational angle is identified by $\theta_2$ that rotates in $x_1 z_1$ plane. Furthermore, the joint $J_3$, $J_4$ represent the elbow and the wrist respectively, making the elbow and the wrist axes of $z$ to be moving in synchronous to $J_2$ that defines the rotational angles $\theta_3$ and $\theta_4$ to be rotated in $x_2 y_2$ and $x_3 y_3$ planes respectively. Last but not least, $J_5$ represents the gripper leading the perpendicularity of the axis $z_5$ on $z_4$ to define the angle $\theta_5$ that rotates in $x_4 y_4$ plane.

7. **Forward Robot Arm Kinematics**

The robot arm sections can be illustrated by specific kinematics denoted firstly as forward kinematics that show the usability of each part in the arm. The interested part of this section is the end robot that can usually be constructed from a gripper and a camera as our proposed sample example. The kinematic describes the mobility and the interaction of each joint and link in the arm separately. It is worth mentioning that the links are connected to each other as chains in order to activate the immediate interaction during the motion. Furthermore, as mentioned before that each section is modeled symbolically to express the number of joints and links, such that the joint $J_i$ represents the current joint, $J_{i-1}$ represents the previous joint, the links $L_0$ to $L_n$ represent the number of links available. The forward kinematic states the fact that robot arm considers the locations $x, y, z$ axes of the object to specify the rotational angles $\theta$, which means that this method cannot help in this proposed work. Hereby, the suitable solution is the inverse robot arm kinematics [19].
8. **Inverse Robot Arm Kinematics**

The rotation and activation procedure of the inverse kinematic differs than the forward due to the drastic dependence on considering the rotational angles to realize the location of a target in \(x, y, z\) planes. Hence, this method is more appropriated for the proposed work scope. The inverse kinematic analyzes and discusses the motion of the robot arm and the rotational technique of the joints represented by angular models. The directions and the locations are interpreted in Cartesian space in order to represent \(x, y, z\) with the corresponding joints angles. The solution offered by the inverse kinematic is not only a sole idea due to the complexity of the technique that deals with various parameters. The technique takes into consideration the interested area of the work by integrating the workspace as a part of robot arm environment. Furthermore, the shape and the size of the workspace can be chosen based on robot arm size and dimension. According to the base joint angle, Figure 5 presents that the robot arm awards the radial distance \(r\) with respect to \(x, y\) plane [19].

![Figure 5. Links and Joints Analysis - Lynx Arm](image)

From top side, an alternative angle identified by \(\theta_1\) is created to represent the angular motion between \(x, y\) axes with respect to the distance as follows [15]:

\[
r = \sqrt{x^2 + y^2}
\]

\[
\theta_1 = \tan^{-1}(x, y)
\]  

The angles \(\theta_2\) and \(\theta_3\) can be calculated depending on the distance allowed between \(x, y\) points from the original point \(o\) and the height \(z'\). While, \(\theta_3\) is calculated based on the triangle created using the lines \(L_1\) and \(L_2\) as following expressions:

\[
z' = z - b
\]

\[
\theta_3 = \frac{r^2 + z'i^2 - L_1^2 - L_2^2}{2L_1L_2}
\]

Furthermore, the angle \(\theta_2\) can be expressed partially based on the angles \(\Psi\) and \(\beta\) as pointed in Figure 5. Hereby, these partial parameters are given as follows:

\[
\beta = \tan^{-1}(z', r)
\]

\[
\Psi = \cos^{-1}\left(\frac{r^2 + z'i^2 - L_1^2 - L_2^2}{2L_1\sqrt{r^2 + z'i^2}}\right)
\]

\[
\theta_2 = \beta \pm \Psi
\]

Where, the plus and minus signs are chosen depending on the value of \(\theta_3\) such that the \(\theta_3\) grater or lesser than zero results the minus and plus signs respectively. Finally, \(\theta_4\) is given based on \(\theta_2\) and \(\theta_3\) as follows:

\[
\theta_4 = \Phi - \theta_3 - \theta_2
\]
9. Results and Discussions

The results are built depending on two possibilities including the EIH robot arm that tracks and grasps a stationary object and a moving object, in the scope of the camera for both cases. In this paper, the motion of the object was analyzed such that the speed and the direction of the object were studied. The proposed target that is detected, tracked and grasped respectively is shown in Figure 6, which shows that the object moves in circular direction and the EIH robot arm adapts its motion, speed and the grasping strategy based on behavior of the object.

![Figure 6. The Circular Motion in Reality](image)

For simplicity, the scenario that is followed by the robot arm to grasp a target was illustrated in details as shown in the flowchart in Figure 7.

![Figure 7. Object Grasping Scenario](image)

It is intended to demonstrate the studies performed on the object followed by the action of the arm as prescribed case studies.

A. Case Study A
In this case the target is a moving object with a speed of 0.5 cm/sec moves along x, y axes. Hereby, the MATLAB code has given the result in Figure 8, which confirms that estimated time needed to pick the object up, is specified by 1.4 sec.
B. Case Study B
The interested result was reached in this attempt based on a moving object with a speed of 1 cm/sec, moves only along y axis. The curve in Figure 9 shows that the time required to grasp the object is specified by 1.19 sec. As an important fact, it can be stated that the speed of the object effects drastically on robot arm grasping time.

In order to confirm the fact reached before, it is intended to extend the case studies to comprise more grasping attempts as demonstrated in the following subsection:

C. Case Study C
This subsection discusses and presents several results depending on the speed of the object with respect to its movement direction. It is worth mentioning that the target is chosen to be a moving object along x, y plane in the entire experiments. However, the speed of the object is different in each experiment specified by (2, 3, and 4) cm/sec. Hereby, the resultant responses are shown in the Figures 10, 11, and 12.
The alteration in object velocity based on (2, 3, 4) cm/sec, resulted grasping time around (2.2, 2.4, 2.9) sec, respectively as demonstrated in the Figures 10, 11, and 12. This means, that the speed of the object is directly proportional with the grasping time and vice versa.

D. Case Study D
The final result can be predicated easily specially if has been known that the target is a stationary object located in a specific location. The response shown in Figures 13 demonstrates the minimum time needed to grasp the object which specified by 0.9 sec due to stability of the object in a single location.
It has to be mentioned that the robot arm can move in both x, y axes in order to be adapted with object motion. Hereby, the entire figures in the case studies display the motion of the arm along x and y axis. Finally, the entire experiments results are demonstrated uniformly as shown in Table 1.

Table 1: The Proposed Experiments Results

<table>
<thead>
<tr>
<th>Case Study</th>
<th>Object Speed cm / sec</th>
<th>Object Direction</th>
<th>Grasping Time sec</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.5</td>
<td>x, y</td>
<td>1.4</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>y</td>
<td>1.19</td>
</tr>
<tr>
<td>C</td>
<td>2</td>
<td>x, y</td>
<td>2.2</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>x, y</td>
<td>2.4</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>x, y</td>
<td>2.9</td>
</tr>
<tr>
<td>D</td>
<td>0</td>
<td>Stationary</td>
<td>0.9</td>
</tr>
</tbody>
</table>

10. Conclusions

In this paper, a unique real time object tracking and grasping technique is proposed based on Eye in Hand robot arm configuration. The technique works to track the target object using Kalman filter method that defines the prediction and the refined stages. Furthermore, the image processing operation is applied to extract the required information of an image captured by the camera to detect object location. The employed robot arm based Eye in Hand (EIH) camera structure provides valuable features identified as: the mounted camera at the end of the arm guarantees motion freedom but with limited workspace coverage area, the preparation for environment sizing is not required, since the workspace is predictable, depending on the first two advantages, the extended random motion of the object can be dominated easily. Kalman filter method is used to define the path of the moving object with respect to different object speeds and considerations. For more clarity, it is intended to demonstrate the tracking and grasping results as follows:

1. The robotic system was tested on moving object with different speeds specified by 0.5 cm/sec along x, y axes, and 1 cm/sec along y axis. Hereby, the simulation results estimated the grasping time of the object by 1.4 sec, and 1.19 sec respectively.
2. For more confirmation, additional object speeds specified by (2, 3, 4) cm/sec are applied with respect to a constant movement direction along x, y axes. Under this act, the simulation results defined the grasping time by (2.2, 2.4, 2.9) sec respectively.
3. Furthermore, the grasping time of the arm was recorded by 0.9 sec for static object probability.
4. The presented results might be enhanced for a little bit more by considering the use of a huge robot arm with less oscillation to magnify the workspace and accurate the prediction.
5. It can be stated that the speed of the object is directly proportional with the grasping time and vice versa, regardless object direction.
6. Finally, as stated in the kinematics that the proposed robot arm is constructed of 1 gripper and 5 DOF joints which forced to apply experiments with limited object speeds.
References


