ZIBELINE INTERNATIONAL PUBLISHING

ISSN: 2222-7059 (Print)
EISSN: 2222-7067 (Online)

Advances In Industrial Engineering And Management (AIEM)

DOI: http://doi.org/10.7508/aiem.01.2020.19.30

AIEM ADVANCES IN INDUSTRIAL ENGINEERING AND MANAGEMENT

CrossMark

ARTICLE

SURVEY OF THE CURRENT PRACTICES AND CHALLENGES FOR VISION SYSTEMS IN INDUSTRIAL ROBOTIC GRASPING AND ASSEMBLY APPLICATIONS

Hamza Alzarok[1*], Simon Fletcher[2], Andrew. P. Longstaff[2]
[1]Department of Electrical and Electronic Engineering / Faculty of Engineering, Bani Walid University, Libya
[2]Centre for Precision Technologies, School of Computing and Engineering, University of Huddersfield, Queensgate, Huddersfield HD1 3DH, England
****Corresponding Author E-mail: Hamza.Alzarok@bwu.edu.ly, Hamza.Alzarok@gmail.com***

ARTICLE DETAILS

Article History:

Received 12 March 2020
Accepted 15 April 2020
Available online 12 May 2020

ABSTRACT

Vision systems have now been involved to a greater extent in robotic manufacturing applications due to their non-contact features and high accuracy. The visual information extracted from the vision sensor significantly assists robot manipulators to perform a variety of tasks with an accuracy that satisfactorily meets the industrial demands. A successful accomplishment of these tasks is heavily dependent on the feedback from the vision sensors to enhance the efficiency of detection, tracking and control of the robot motion by utilising their visual information. The feedback, therefore, enhances the safety of the system by preventing the robots from being damaged and operators from being injured which, in turn, saves the production time. In the modern industry, there is an increasing requirement for advanced robot-based target detection and tracking, target grasping and also for the capability to execute assembling tasks in unprepared environments with randomly positioned/oriented targets. Grasping and Assembly tasks represent the most important applications for industrial robots that often require the additional feature of vision systems as a navigation guidance for tracking and intercepting of moving targets. This paper targets these application areas and presents a review of the state-of-the-art equipment, methodologies and practices used within the associated research areas of robotic systems in the context of vision systems. It also examines the recent contributions of the vision systems in robotic tasks and highlights on their performance, the use of algorithms for image processing and calibration procedures adopted, and their contribution towards the effectiveness of robotic positioning resolution and accuracy.

KEYWORDS

Vision Systems, Industrial Robots, Machine Vision, Object Grasping, Pick-and-place Tasks, Robotic Assembly.

1. INTRODUCTION

An industrial robot is defined by the International Standards Organisation (ISO) as “an automatically controlled, reprogrammable, multipurpose manipulator programmable in three or more axes” [1]. The field of robotics may be more practically defined as the study, design and use of robot systems for manufacturing. The first industrial robot was invented by George Charles Devol who is often called the father of robotics [2]. Traditionally, they have the capability to change their position repeatedly with a small repeatability error in the sub-millimetre range of ±0.3 mm, but their absolute accuracy can be in the order several mms (±5–10 mm, ±0.5–1.8 mm) because of mechanical tolerances, elasticities, temperature, etc [3-5]. These error sources can cause a significant offset to the robot end-effector. Therefore, it is important for applications that require an absolute accuracy to measure the end-effector position and orientation in Cartesian space [6].

The advantage of using the articulated robots in the industry is because of their freedom of movement in the environment and also for their ability to perform tasks without being fixed to one physical location. In fact, there is a high demand for using mobile robots in unprepared environments and hard-to-reach or hazardous areas [7]. In such scenarios, robots are often operated from a safe distance. In order to ensure an error-free performance of mobile robots, monitoring systems must be installed to keep tracking of the motion parameters programmed within mobile robots. The performance of mobile robots can be further enhanced by using the feedback from the installed sensors which can help robots to perform complex tasks by acquiring information such as, the robot location. In the past research, reviews have been conducted which highlighted the sensory techniques used for robots such as provided a good review on the fundamentals of inertial and visual sensors which was followed by an intensive review from about the sensor systems that can be used in a specific kind of robot (flexible manipulators) [8,9]. However, the research or literature for the applications of sensory techniques within the industrial robots area does not address the current practices of vision sensors used within the areas of both applied and research oriented solutions for monitoring or improving industrial robot performance which has become the main motivation for this article. This paper provides a comprehensive review of the current contributions of these systems for detecting and tracking of industrial robots and also the challenges toward achieving the demanded manufacturing performance.

There are two possible solutions for measuring the robot position. The first solution is to use embedded sensors for measuring the position of the articulated robot, such as by using accelerometers [10] and joint sensors [11,12] for measuring the angular position and velocity as well as the Cartesian position, orientation, and the linear velocity of the robot end-effector. The advantage of these methods is the ability to provide good short term position estimates. However, they suffer from errors accumulating over time because of the integration of minute increments measurements (e.g. inertial) to obtain the final estimate or from errors outside of the control loop such as deformation of the links either from time varying or finite stiffness effects [13]. The second solution is to measure the robot position globally using external sensors such as laser trackers, vision systems or and intelligent global positioning systems (iGPS) [14].

Large metrology systems such as laser trackers and iGPS provide higher positioning accuracy than other sensors, and have the capability to be used for measuring the position of the robot end-effector in small and large measurement volumes. However, in order to achieve a sub-millimetre accuracy by these sensors, the hardware cost will increase. Table 1 shows the advantages and drawbacks for various sensory techniques used for measuring and tracking the end-effector of the articulated robots.

The Global Positioning Systems (GPS) is one of popular sensors that have been used for tracking of articulated manipulators, they provide a sufficient and accurate way to track robots, but it is not applicable indoors, therefore a recent innovation in GPS is named indoor GPS (iGPS) can be used indoors for industrial robot tracking with a positioning accuracy in the sub-millimetre range if it was setup with at least four transmitters [7,15,16]. However, the iGPS is more suited for large measurement volumes. This is due to the need to set up the transmitters which must be visible to the sensor during the tracking process. In the cases where only three of four transmitters are visible, the accuracy then will be reduced by 10%–15% of the overall accuracy [17].

Laser trackers and photogrammetry are also typically used in dynamic tracking measurements. Although these systems provide high sufficient performance, they are subject to two main limitations which are: 1) the number of points that can be measured simultaneously by the laser tracker and 2) the measurement area for laser tracker and photogrammetry. Additional systems must be purchased in order to expand the coverage in the area of measurement and/or number of points which rise the cost of the tracking system [14].

Optoelectronic position sensors have been also used by researchers for detecting and measuring the position and orientation of a moving articulated robot [18], and also for estimating its motion speed by fusing the information of the sensor with accelerometers and gyroscopes [19]. The reason behind applying a PSD instead of CCD/CMOS sensors is the ability to sense the position of a light spot on the detector which makes it more appropriate for real time feedback applications and also a faster response time and higher sampling rate can be achieved compared with the other sensors. However, the PSD sensors cannot capture images as vision cameras do and that is the main weak point (i.e. its incapability for sensing without markers). Moreover, they are more fitted to small measurement volumes than large measurement areas which require a multi-camera system.

Low cost depth sensors such as The Microsoft Kinect sensor [20,21] were also used by researchers for estimating the position of articulated robots in 3D space. The working principle of the sensor relies on the integration of 3D range and colour information in order to obtain real-time 3D tracking data. However, it was designed and used for human body motion tracking in 3D space with an accuracy of few millimetres. Moreover, it is not suited for high precision tasks because of its low accuracy and also because of the reduction in its tracking performance due to the occlusion problem that causes adding fake colour information to some pixels in the depth colour images.

A fringe projection device (namely IScan M300 by Imetrics) was also used by researchers [22] for estimating the position of robot end effectors. The fringe projection is often applied in dimensional metrology, especially in Quality Control and Reverse Engineering applications. It has the capability in performing trajectory tracking with high positioning accuracy (i.e. the highest positioning error calculated along the selected trajectory was 43 µm for a task demanding an accuracy of ±100 µm), in addition to that, it does not need alignment devices, and also provides 3D measurements because of the combined translation and rotation. However, the main drawback of this measuring system is in terms of time consumption. Due to the limitation of its FOV, this device can only be used for a small measurement volume, and in order to track a large number of points along the motion path of the robot end-effector, there is a necessity to split the measurement area into smaller areas and keep the position and orientation of the robot end-effector stationary during the relocation of the equipment from one area to the next.

In order to balance between the hardware cost of the measurement system and its capability to provide sufficient information about the robot, vision sensors have been introduced, the distinguishing point of these sensors is in the reduced hardware cost, the easiness to install and use, and also for their applicability in both small and large measurement volumes. In industrial robotic applications, vision sensors have be intensively employed with the robots as tracking systems, research trends of using these systems have significantly increased in the last two decades (as shown in Figure 1), particularly in popular automation applications such as pick-place (or grasping), and assembly. Moreover, due to the shortage of well-trained labourers and the minimising manufacturing prices of products, many industrial companies are focusing more on automation in their production processes. Therefore, it is expected that with the rapid developments in the vision technology, the automotive industry will hold the largest market in the near future. According to a market research report released by Markets And Markets Analysis [23], the global market size for machine visions is expected to grow from USD 10.7 billion in 2020 to USD 14.7 billion by 2025 (see Figure 2). The report stated that the growing demand for vision-guided robotic systems is the major driving factor in this market.

In this paper, the contribution and challenges of the vision sensors in automation applications, particularly in robotic grasping and assembly tasks will be our scope. Also, this paper will highlight on the efforts made by researcher to cope with the visual tracking problems. The objectives of this work can be listed as follows: 1) highlights on the types of camera configurations, their uses in different tracking scenarios, and their advantages and drawbacks, 2) review for the uses of vision sensors in robotic grasping and assembly tasks, the review will provide a description of the camera hardware and robot specifications, type of image processing algorithms been used along with vision sensors, the description of the object target to be tracked (shape and appearance representations), the measurement volume and the environment, and 3) highlights on the different limitations that restrict the use of vision sensors in the two aforementioned tasks, also highlights on both hardware and software solutions that have been introduced by researchers to cope with tracking problems.

This paper is organized as follows. Section 1.1 highlights on the possible configurations of the vision sensor, Section 1.2 reviews the use of vision systems for guiding the robots for pre-grasping (intercepting) and grasping of targets. Section 1.3 reviews the use of vision systems in robotic assembly. Section 1.4 discusses the associated challenges which may affect the applications of the vision system in the industrial robotic fields followed by Section 1.5 which outlines the conclusions drawn from this work.

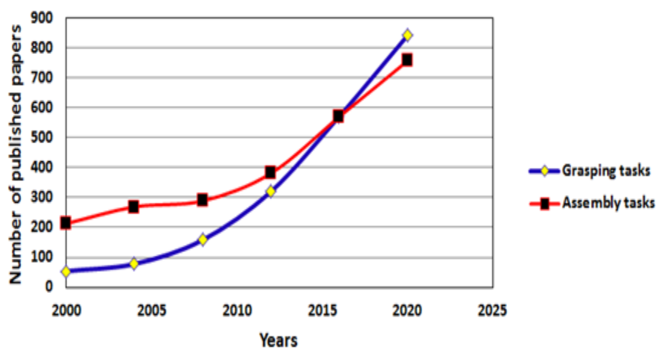

Figure 1: Trends in the use of vision system in robotic assembly and grasping tasks in the last two decades.

Table 1: Some typical external measurement systems capabilities and limitations.

Type of sensory technique	Citation	Advantages	Drawbacks
Indoor (iGPS)	[15-17, 24-28]	Positioning accuracy is less than 0.5 mm. comparable to a laser tracker at low speeds of 100 mm/s	• The achieved accuracy is subject to many factors such as the number and layout of visible transmitters, environmental conditions, sampling duration and sensor velocity and the synchronisation between the tracker and robot • The setup is more suited for large scale measurement. • Expensive hardware system ($150,000).
Laser tracker	[14]	Positioning accuracy is less than 0.5 mm.	• Additional systems must be purchased in order to cover the volume of measurement which increase the cost of the tracking system. • Expensive hardware system ($100,000).
PSD sensors	[18, 19, 29]	PSD sensors require less time for image processing compared to CCD sensors. It is preferred in a certain environment when the intensity of the target is lighter than its environment due to their ability to quickly and directly transduce the projected position of the light into an analogue current. Can track at high robot speed (1m/s) with a low hardware cost (around $450)	• incapability for sensing without markers and no typical accuracy reported.
Microsoft Kinect sensor	[20, 21, 30]	Low hardware cost (around $110)	•Positioning accuracy of few millimetres tracking performance of the sensor reduces due to the occlusion problem.
Light scanners	[22]	The highest positioning error calculated along the selected trajectory was 43 μm for a task requiring an accuracy of ±100 μm	• The drawback is in terms of time consumption. Also, there is a necessity to divide the measurement area into smaller areas and keep the robot stationary during the relocation of the equipment from one area to the next.
CCD/CMOS	[31]	Low hardware cost (between $1000 and $2000)	• No typical accuracy reported.

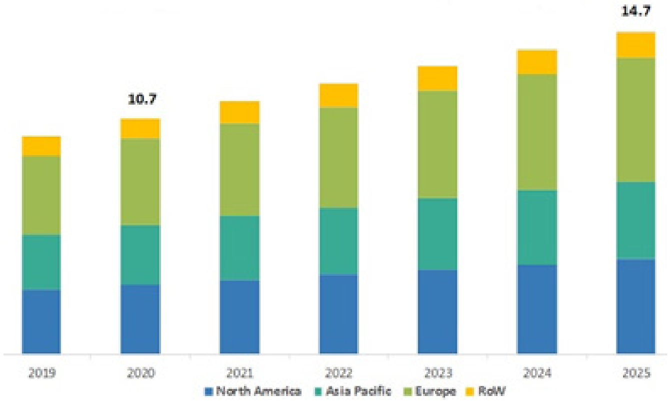

Figure 2: Machine vision market, by region (USD billion). Note: RoW includes South America, Middle East and Africa [23].

1.1 Vision sensor setups

Visual tracking can be defined as "the process of locating, identifying, and determining the dynamic configuration of one or many moving (possibly deformable) objects (or parts of objects) in each frame of one or several cameras" [32-34]. The simple definition for the visual tracking was given by [35] as "Controlling the pose of mobile robot using a vision system to track a target". In this section, a brief review will be performed on the use of the camera setups for tracking tasks in industrial robotic applications.

In a multi-arm robotic cell, there are two possible configurations to set up a vision system (see Figure 3) namely eye-in-hand and eye-to-hand [36-42]. The eye-in-hand camera configuration (ENH) involves visual systems that are usually used by composing two or more cameras that can be rigidly attached to the robot end-effectors, whereas the eye-to-hand camera configuration (E2H) involves the vision systems that are fixed in the workspace [43]. The ENH camera method enables the robot to view the workspace more flexibly, but with limited field of view (FOV)

[36]. However, E2H camera ensures a panoramic view of the workspace, but typically with a lower accuracy [44].

As the single ENH camera moves away from the target, the FOV of the camera is increased at the cost of reduced accuracy. Multi ENH configurations with different FOVs were proposed to cope with deficiencies of a single camera and to increase the total FOV and improve the overall accuracy. However, the visibility of the target object, and the common existence of errors in the target object modelling (e.g., CAD modelling errors) are crucial for such configurations and strongly affects the overall robot pose estimation. According to there are a few methodologies introduced for addressing this problem such as one by in which comparison was made between the precision of multi ENH camera to a single ENH camera. The authors found that the positioning accuracy was significantly improved with the use of multi-camera visual Servoing compared to the use of a single camera [45,46].

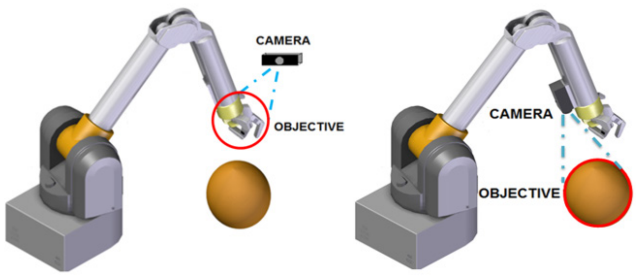

Figure 3: Possible camera configurations.

Another case which is less investigated in the reviewed papers is the cooperation between two cameras, the first one could be considered as an E2H camera, and the other one as an ENH camera see [47,48]. The advantage of this cooperation is the ability to subdivide the tracking process into two tasks, the first is tracking the target and the second is knowing the position of the robot end-effector, with each task being performed with a single camera. In both tasks, image information is used to measure the error between the current location of the robot and its desired location. The image information that is used to perform the task is either i) two dimensional, described by using coordinates of the image plane, or ii) three dimensional where a camera or object model is employed to retrieve pose information with respect to the camera/world/robot coordinate system. The robot is controlled either using image information as two or three dimensional which classifies the visual servo systems additionally as: Position-based visual servo systems, Image-based visual servo systems, and the third is a cooperation of the previous two approaches, called 2 ½ D visual Servoing [49]. This approach provides an advantage in not requiring any geometric 3D model of the object in comparison with position-based visual Servoing. Furthermore, the 2 ½ D visual Servoing ensures the convergence of the control law in the whole task space which is not available in the image-based visual Servoing. However, the 2 ½ D visual Servoing is more sensitive to image noise in comparison with 2D visual Servoing, and that refers to the third approach which directly uses visual features as inputs of the control law without any additional estimation step [50].

An ENH/E2H camera cooperation is introduced as suitable choice in the case where the employed stereo vision system either cannot provide precise information or can provide accurate information about the actual position of the robot but with long time delay. In this case, each task is performed with a single camera i.e. the ENH camera can perform the tracking of the target and the fixed camera can measure the position of the robot arm in the 3D space. Moreover, while complex artificial vision algorithms are required by the stereo vision system, an ENH/E2H camera cooperation could be performed with easier image processing [51].

In order to obtain additional visual information about the tracked targets, the use of multiple cameras is often a more suitable choice compared to single or stereo camera configurations [52]. However, the drawback of using multi-camera systems is the need for matching across multiple views that are captured from different cameras having different perspectives, which is usually a time consuming and non–trivial problem. Therefore, servo systems that use more than two cameras for controlling a robot are uncommon, in addition to the matching problem, and the additional hardware cost as main factors affecting the use of multi-cameras [49].

In the next sections, the use of ENH and E2H systems as tracking systems with the robotic manipulation will be highlighted, then the advantages and disadvantage of each system will be listed at the end of each section.

1.2 ENH setup

Cameras in an ENH configuration had been widely used by researchers in tasks that require detecting and tracking the location of the target to be manipulated, they are amounted on the end effector of the robot, therefore, this kind of setup cannot be used for detecting and tracking the position of the robot end-effector. One of examples for the use of ENH setup in tracking tasks was introduced by [53]. The ENH setup is used to track three particular circular features of the object (an engine cover) at 60 frames per second (30 fps per camera). According to the authors, the proposed tracking algorithm provides very good accuracy compared with other algorithms that are based on CAD models or unconstructed cloud of points [54, 55]. Due to the difficulty of measuring the error for a moving object, the object was kept static while the camera/end-effector was moving along a random path. The proposed algorithm is only suitable for use in controlled environment such as consistent lighting conditions. Results obtained showed that there is no dependency between the uncertainty in the position of the features and the distance between the camera and the object. These unexpected results were due to some factors which have an effect on the estimation of the pose, these were the condition of the illumination, and the tracking quality for different sizes of the object features as perceived in the image.

Jia also used an ENH system which combines a monocular camera fixed on the tip of a 6-DOF robot [56]. The target object was represented as a set of points while iterative learning control was used to provide visual trajectory tracking. Only images with resolution of 0.3 megapixels were used for recording the trajectory tracking and the desired velocity trajectory was only used for analysing pose error based on kinematics. The results showed that the convergence of the pose error was minimised to less than 0.5 mm after 5 iterations. However, due to the presence of image noise, the tracking error cannot converge to zero.

An ENH system was also addressed by Hua for 3D target tracking [57]. The robotic system was composed of a 6-DOF robot and a camera fixed at the end of the robot. The sliding mode control and adaptive technique were employed in the design of the proposed controller in order to remove the nonlinear function of the robot. The effectiveness of the proposed controller had been proved in simulation and on a consideration of using a two link robot manipulator, the results showed that tracking errors between the robot and the moving target in word coordinates approximately converges to zero after around 5 seconds of starting the motion of the target.

The advantages: the location of the camera is always stationary relative to the hand of the robot. The target object to be captured and detected cannot be hidden from the camera by interference from the robot manipulator [58].

The drawbacks: the ENH cameras are so bulky that even the smallest sized camera is too large to allow appropriate physical integration with the manipulator's gripper. Thus the manoeuvrability of the manipulator system is constrained [58].

1.3 E2H setup

Cameras in an E2H configuration had been widely used by researchers in tasks that require detecting and tracking the position of the robot end-effector. One of examples for the use of E2H setup in tracking tasks was introduced by Wang, in their work, a new controller was proposed for controlling feature points on the end-effector of the robot to trace desired trajectories (circular trajectory) defined on the image plane of the camera, the proposed work was an extension of their previous work [59,60]. The main difference between the previous and current is that in their old work, the controller can only guarantee semi-global stability, however in the proposed work, the developed controller can provide a global stability of the system. Another advantage of the new algorithm is in its ability to cope with the unknown feature positions. The performance of the presented controller for trajectory tracking of features on a 3-DOF robot was demonstrated only by the simulations in their work. It should be noted that the measurements of the position errors (the difference between the desired and actual trajectories) were simulated in image coordinates (pixels) instead of the world coordinates (e.g. mm).

Chang used an E2H for another robotic task, the task required the use of the camera to drive the spot of the laser in order to follow a visually arbitrary planar contour [61]. Only two of revolute joints of the robot were used to perform the 2-DOF rotational motion (pitch and yaw motion). A laser pointer was mounted on the end-effector of the robot whereas the camera was used for tracking the spot of the laser and geometric features of the laser pointer (i.e., the centre). The reference trajectory of the vision system was obtained from the fixed planar contour. Moreover, in order to detect the centre line of the laser pointer, three marks (dark strip features) were made on the cylinder-type laser. In the performed experiments, the diameter of the laser spot was about of 12 pixels, and the tracking error measured was less than 4 pixels.

Wang used a vision system (E2H) which is mounted near a 3-DOF robot (Puma 560) to monitor its motion with the camera providing a video signal [62]. The target object was set as a one feature point on the robot end-effector and was extracted from the images. The robot was programmed to move in two different trajectories (linear and circular). An adaptive dynamic controller was proposed for the image-based trajectory tracking of a feature point on a robot manipulator in an un-calibrated E2H setup. Simulation results showed that the tracking errors on the image plane for both trajectories had been significantly reduced along the tracking process time. However, the performance of the proposed controller could not be easily evaluated due to that the results of the trajectory tracking experiment were only available in simulations. The ENH setup was proposed for future work in order to extend the applications of the proposed technique.

A multi E2H system was used by Grosso for tracking of the robot end-effector trajectory during the execution of reaching tasks [63]. They aimed in their work to make the robot reaching the target by applying continues correction for its trajectory in order to react to the cap motion. An optical flow was used to obtain visual measurements that are used for controlling the robot. In order to cope with the computational speed of the tracking system, the original image was sampled down to 80×80 pixels and limiting the computation of the optical flow was suggested as a further enhancement. Moreover, the position of the cameras was adjusted for each test in order to cover the FOV of the working area.

Kragic also dealt with the problem of the computational speed for the tracking system and the FOV for the camera [64]. In order to cope with the first challenge, they introduced a template matching algorithm which is initialized in the region where the area of interest was found in the previous frame, and to deal with the second challenge, two colour CCD cameras with a controllable pan–tilt unit was used for keeping the robot end–effector centred in the image. However, no mention was made in their presented work whether the accuracy of the pan-tilt camera has any effect on the tracking performance, although the camera had low resolution and its view to the measurement volume was from a distance of about 2 m.

The use of a multi camera system or a stereo head camera with a controllable pan and tilt unit are not the only suggestions by researchers for coping with the restriction of FOV during the use of E2H camera setup. For instance, Chien used an E2H system which was fixed above the workspace in order to obtain a full view of the robot workspace [65]. The aim of their work was to move the 2-DOF planar robot end effector to follow a desired circular trajectory in the Cartesian space via the observation from the camera located at the end effector position in the image space. The Function Approximation Technique (FAT) was proposed to deal with time uncertainties.

Lund also used a one CCD camera mounted under the ceiling of the lab about 2 m above the measurement area in order to obtain a full view of the moving robot, in their work, they also suggested using a small observation window around the estimated position of the robot and keeping the speed of manipulator at approximately of 1 m/sec for decreasing the amount of processing data during the tracking process [66]. The drawback of their system is in their use of LEDs for detecting the position of the robot where there is an absolute dependency of the tracking system on them.

Gupta also used a camera mounted on ceiling of the workspace, a panoramic camera was used rather than other cameras because of its ability to provide a wide view in the entire room (i.e., 180 degrees of FOV), making the analysis of the projection much easier compared to that of using a multi camera system [67]. However, in their work, they reduce the resolution of the captured images in order to increase the image processing speed, and a Kalman filter was used to enhance the tracking accuracy. Although the results showed the effectiveness of the proposed method for tracking and detecting the pose of the robot in indoor environment. Two problems were faced during the tracking experiment; the first is the inability of detecting the position of the markers in the camera image, and the second is the incorrect detection for the markers in the camera image. The first problem can be solved by stopping the robot from the motion. However, the second problem cannot be easily detected and its occurrence is rare according to the author. The tracking process can fail due to the light sources in the room, occlusion of the robot markers or in the case of covered lenses of the camera.

The restriction on the motion speed of the robot is not the only challenge related to the robot, the variation in its motion scenario is another challenge. Baek mentioned the difficulty of estimating the next position of the robot because of its movement which can be varied with different types of motions (i.e., changing the motion pattern with large rotation angles leads to inability of the algorithm to correctly predict the end-effector position) [68]. They suggested Kalman filtering techniques as typical solutions for this problem.

The weight of the robot is another issue which is mentioned by Wang, they noted a strong effect of nonlinear forces on the motion of the robot, since the robot was light [69]. The effect was caused by the small values for the gear ratios and input motor power. Their work aimed to track the trajectory of the robot end-effector using a Hough transform which identifies point and line features of the tracked target. However, there was no indication in their paper on how to cope with the computational cost of the Hough transform.

A cooperation of calibrated ENH and E2H cameras was used by Elena for monitoring a moving target and measuring the position of the robot-end-effector in the 3D space [51]. In their work, they suggested using adaptive controllers with a variable structure in order to cope with the changes in the dynamics of the target. In the case of using a multi camera system for trajectory tracking tasks, the correspondence information between the two cameras about the required trajectory is necessary, therefore, a new strategy based on implementation of the path tracking control was presented by Chang in the case of the absence of these information [70,71].

Advantages: E2H camera ensures a panoramic view of the workspace, but typically with a lower accuracy [44].

Disadvantage: the field of view of the E2H system is obstructed while the robot manipulator is retrieving the object. Each single E2H will be focused on only a part of the working area, with no possibility to zoom in on that area with increased resolution [58].

1.4 Robotic interception and grasping tasks

Robotic interception is an important task for industrial robots and which typically requires vision systems as navigation guidance systems for tracking and intercepting of moving targets [72]. This task can be defined as a pre-grasping stage in which the end-effector of the robot is brought to a pre-grasp rendezvous location with the moving target in the shortest possible time (see Figure 4). The interception of the moving target using an industrial robotic system can be classified based on the type of the target's motion into 1) an interception of manoeuvring objects and 2) an interception of non-manoeuvring objects.

A target can be called as "manoeuvring" if it changes its motion randomly and quickly which causes a difficulty in achieving time-optimal interception. Whereas a target can be described as "non-manoeuvring" if it changes its motion on a continuous path with constant speed or acceleration hence allowing accurate long-term prediction of the motion of the target for time-optimal interception [73, 74].

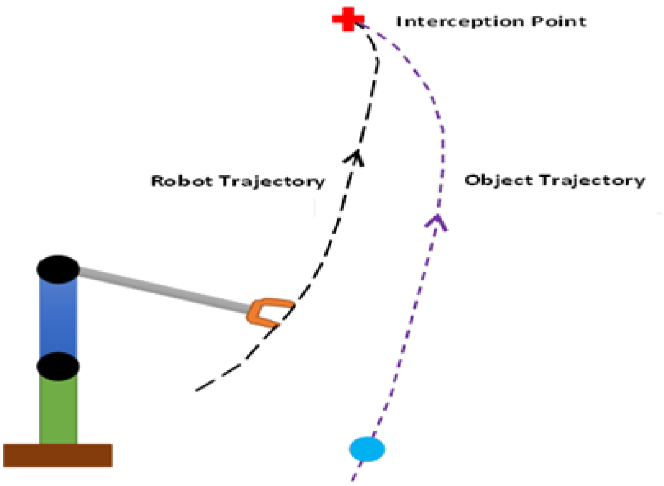

Figure 4: Interception of a moving object using the robot.

Buttazzo used a fixed single camera (E2H) for guiding the robot to grasp a moving object. Simple centroid computation algorithm based was used for detecting the object motion, whereas a constant acceleration assumption was used for predicting the next motion and compensate delays occurred by the vision sensor [75]. Real-time interception of manoeuvring and non-manoeuvring objects was addressed by Borg, their work showed the affectivity of combining navigation technique (IPNG) with a tracking technique (QPT) in reducing the interception time [73].

Grasping a target moving in a complex path such as sinusoidal path is a challenging issue for vision systems, work introduced by Keshmiri showed that complex motion paths require extra robot motion which will led to unnecessary following of the target by the sensor, and thus delay in grasping time and energy consumption [76]. ANN was suggested further to cope with the grasping time delays. Moreover, an optimum grip or a stable grasping for the target is another challenge for vision tracking system, a strategy such introduced by Fuentes-Pacheco [77] which relies on using a multi E2H camera system and a linear prediction method cannot guarantee achieving that, they suggest employing proximity sensors in the robotic end-effector.

The presence of the noise in the vision data is another issue affecting the ability of the sensor in obtaining stable estimation of the grasping time, Nomura used a hybrid Kalman filter for obtaining stable estimation against the noise, the filter considerably reduces the image error (from 25 pixels to 2.0 pixels) during a slow motion scenario (at 0.03 m/s) [78]. However, the estimation process for the grasping point cannot be easily performed in the case of high motion scenarios (i.e. the speed of the robot and the target to be grasped are moving at 1.5 m/s), therefore, it is important to use a proper detection method such as introduced by Gupta who used a blob detection technique for tracking moving balls with colour patterns [79,80], and a STBC (State Transition Based Control) algorithm for implementing the robot behaviour based on the data obtained from the vision processor. The detection technique should have the capability to identify the target to be grasped a mong different objects in a pick-and-place task, such as the one introduced by Shaw [81] which is a combination of SURF (Speed-Up Robust Feature) algorithm with back projection method.

Ge revealed two main challenges in the visual servoing for robotic grasping tasks which are 1) solving the delay caused by image processing or the response of the robot system and 2) resolving the occlusion of the target [82]. These shortcomings are the major reason for a limited performance in the tracking and grasping process which can be solved through the use of predictive algorithms [77].

Gengenbach dealt with the occlusion problem by employing an Extended Kalman filter, and in order to avoid collisions, the common field of view of their multi cameras were out of the workspace of the robot [83]. Moreover, in order avoid the computational cost, the lens distortion was neglected. However, the authors have not addressed the effect of lens distortion on the feasibility of the grasping object.

The influence of the distortion problem on the object appearance in the captured image data was addressed by Kim who stated other sources for the distortion problem which are illumination reflection, the material of the object surface, and the angle of the camera [84]. They proposed a polarizing filter in order to cope with the distortion problem.

Detecting and tracking of the object to be grasped in every captured frame and providing the information for an industrial robot to grasp is another concern, especially when the object is slowly moving on the conveyor, therefore, Guo used an embedded image card with a FPGA (Field Programming Gate Array) for accelerating the processing speed [85]. The detection algorithm was optimized and implemented on a DSP card for enhancing the real-time performance. The authors focus on developing the performance of the vision system in detecting and tracking of the moving objects before being properly grasped by the robot arm. In other words, tracking process for moving objects is very important in tasks that require a quick action from the robot (a quick interception and grasping), researchers [73,75,78,79] used vision systems for providing an accurate estimation of objet motion, the time required for the robot to intercept with the moving object or execute grasping was used as a standard for evaluating the performance of the tracking system.

In addition to the significance of identifying the object to be manipulated, generating the pick-and-place motion of a known object in real-time requires a successful plan for the trajectory of the robot arms to perform the pick-and-place task [86], therefore the configuration and position for the robot end-effector needed to be identified [87].

In image based visual tracking, there are two problems that need to be solved. The first one is how to recognise and detect the target object in the image plane, and the second is how precisely estimate its relative position without a priori information. Due to the difficulty of segmenting the target object's motion from a background in unprepared environment, robust detection of the target is often performed via artificial markers that are not difficult to segment [35]. Distante introduced a solution to the problem of controlling a manipulator to identify and grasp the object [88]. In their work, they used objects which have spherical shapes and known sizes. Cowan used a cubic target on one surface of which several visible marked symbols have specific configurations [89]. If these symbols were detected, then the pose of the target object could be estimated from the configuration of the symbols. Table 2 shows different simple shapes were selected for identification of the grasping task, some researchers used different detection algorithms based on the utilisation of the object shape features such as in [88] whereas in some works [90], the distinctive colour features was used for identifying and grasping the object. However, in the case of general objects with complex shapes will be used, robust methods of detecting a target are necessary in dynamic environment.

Berscheid dealt with another challenge in the grasping tasks which is how to teach the robot to avoid surrounding obstacles that might not permit any possible grasps of a certain object [91]. In their work, they used a combination of a stereo camera (namely Ensenso) with a Franka robotic arm, a Neural Network algorithm (NN) was used along with the extracted visual information for teaching the robot on shifting and grasping of different shaped objects. The camera was placed on the robot's flange (ENH setup), the reason behind the selection of this setup is to obtain fixed transformations between robot, vision sensor and world coordinates. The results showed a high grasp rate with an average value of 92%±7%.

Grasping of unseen objects in cluttered environments is also challenging problem. The reason behind this problem is due to the dependency of many researchers on the use of 2D visual information rather than 3D in the grasping tasks. Choi [92] stated that the use of 3D visual information will make a perfect utilisation for the full workspace of the robot, and will provide suitable grasps for any object within the robot's work volume. The researchers used a depth camera fixed to the upper torso of the robot, and a 3D deep Convolutional Neural Network (CNN) model was used for grasp prediction. The reason behind the selection of depth cameras refers to two reasons: 1) there is no change in its photometric information such as colour and texture, and 2) it utilises a geometric information closely related to target grasping.

Table 2: Examples for different simple shapes being used as grasping targets.

Authors	The detection algorithm	Object shape
[75]	Simple centroid computation algorithm	A Ping-Pong ball (spherical)
[73]	Quintic polynomial tracker	A white circular marker on a darker background (circular)
[76]	Augmented Images Based Visual Servoing	A cube
[77]	Linear prediction through the method of Burg	Red ball (spherical)
[78]	A hybrid Kalman filtering	Industrial part with four holes (circular)
[79]	A blob detection technique based on color	Balls with color patterns (spherical)
[90]	Disparity color and template matching algorithm	A green coloured mark (rectangular)
[83]	Extended kalman filter +optical flow	Screw (elliptic shape)
[88]	Canny operator + Hough transform	Spherical shapes and known sizes
[81]	SURF algorithm + back projection method	Colored objects with different shapes
[93]	Controllable region of interest based on circular hough transform (CRCHT)	A red Ping-Pong ball (spherical)

1.5 Robotic assembly tasks

Robotic assembly systems have been widely used in a range of manufacturing applications. The advantages of automatic assembly compared with manual assembly operations are the minimisation of unit costs, a consistent high quality, the avoidance of hazardous manual operations and the rise of the production standby capacity.

The peg-in-hole assembly is the most popular and often one of the most significant procedures in large-scale equipment manufacturing, which has a significant impact on the product quality. However, in the process of peg-in-hole assembly operation by using robot, the parts may not be assembled correctly or even be damaged due to the deviation caused by the inaccuracy of the assembly robot. In order to cope with the peg-hole jamming problem between the peg and hole, vision technologies are used for measuring the assembly error in order to minimise the peg-in-hole position and orientation errors caused by robot and then eliminate the peg-hole jamming, such as compliance mechanism [94].

Okumura developed an automated visual inspection system that used two high-speed cameras with stereovision to detect mating errors in a peg-base [95]. The peg was highlighted in the image using a semiconductor laser and the system was able to successfully predict 80% of assembly scenarios (successes and failures).

Huang constructed an economic flexible automatic assembly system with the integration of SCARA robot and a simple CCD camera system [31]. The machine vision was introduced to search the locations and measure the size of the assembly parts in the robotic workspace. The experimental results showed that the goal of this robotic random assembly operation was achieved with the appropriate integration of robotic motion control, machine vision, sensor, and assembly force detecting techniques.

Jayaweera used a non-contact sensing system by testing two perpendicular structured light stripes to determine the centre of the hole and measure the part deformation and misalignment [96]. The experimental results showed that the method proposed could be used for the assembly of compliant aero-structure components within required tolerance limits. Ramachandram used a visual sensor consisting of a camera and a grid pattern projector, trained a neural network learning the complex relationship between the robot's pose displacements and the observed changes in the image features [97]. After training, visual feedback guided an industrial robot to complete the intelligent assembly.

Vision sensors have also been used for recognising and detecting the proper target object, Sahu used an integration of vision sensor with ultrasonic sensor in order to perform the recognition [98]. The vision sensor was used for verifying the correct shape and size of the target and the object was identified via through template matching technique. The data obtained from the ultrasonic sensor is used for calculating the grasping point which makes the robotic assembly system more intelligent.

The sensing system is one of the basic components in precision assembly. Microscopic vision systems are often introduced for measuring the pose of the assembling Micro Electro Mechanical Systems (MEMS) components with different range sizes starting from several to hundreds of microns, Microscopic systems can be classified into three types: monocular Zhen, binocular Li, and multi-ocular systems Zhang [99-101]. According to Liu [102], at least three cameras are required to be used for measuring the pose of a 6DOF component to be assembled, since the use of a single camera can only provide measurements of 1 DOF rotation and 2 DOF translation Xu [103].

The use of two or more microscopic cameras may cause blocking in views when the separate axis is not parallel with the camera, and this problem particularly happens when a platform with several manipulators are used for performing assemblies of multiple objects. Xing [104] dealt with the blocking problem caused by the intersection of two or more objects, they proposed two techniques for detecting collisions between cylindrical objects, the techniques are named reconstruction method and projection approach, the late technique has a low computational cost and is more suited for constant assembly operation.

A plan for collision avoidance in the precision assembly was introduced by Xing, two microscopic cameras were used for determining the collision, and a method of constructing 3-D position and posture based on employing two planar projections was used for deciding collision state [105]. The work only considers convex objects and the objects are all dealt as solid-shaped for collision judgment.

Hsu introduced a developed vision system for guidance of a SCARA robot to pick and place surface mount device components on a PCB board [106]. Hough transform was used for detecting the spherical features of the PCB, the results showed that the auto assembly accuracy was significantly enhanced by 78.66% when the manipulator was operated in a teaching mode, the overall accuracy could be further enhanced by 2% if improvements are made on the hardware structure.

Ahn introduced an automatic off-line teaching method using vision information extracted from an ENH camera for robotic assembly [107]. However, in their work, the setup of their proposed method requires between 1 to 2 hours due to the long-time consumption in the camera-to-robot calibration process, but after completing the setting-up period, the proposed technique significantly minimised the teaching time by around 98%.

Davies presented a system to improve the performance of industrial robots operating in unprepared environments [108]. A cooperation of a stereo vision with an artificial neural network (ANN) are used for recognition of the male and female assembly parts in the manufacturing cell, the object recognition process is based on utilisation of the image histogram and image moments which are fed into the ANN to identify the structure of the component. Their work showed the architecture of vision system is being enhanced for handling complex 3D objects in manufacturing applications.

In some assembly tasks, it is difficult to separate the task that can be executed automatically from the tasks that need manual assembly along an assembly line, therefore, the possibility for combining manual and robotic assembly in one work station was investigated, these tasks require cooperating between the robot and human operator as shown in Figure 5, the collaboration between robots and human operators in manufacturing assembly lines is often called cobots [109-111]. Therefore the use of vision systems is one of the suggested solutions which can be used for stopping the robot from moving into areas where manual assembly works are observed. However, the solution based on the use of vision sensors have not yet been entered into industrial use [112].

Increasing the safety of the shared work cell is a challenging issue, therefore, several works have been introduced to cope with the safety issue by integrating the vision sensor with force sensor, the vision sensor is used to perform two jobs, supervising of the assembly process (for instance monitoring or detecting the tool insertion), and for activating the emergency stop in the case of the worker hand was detected in the scene during robot motion [113-115].

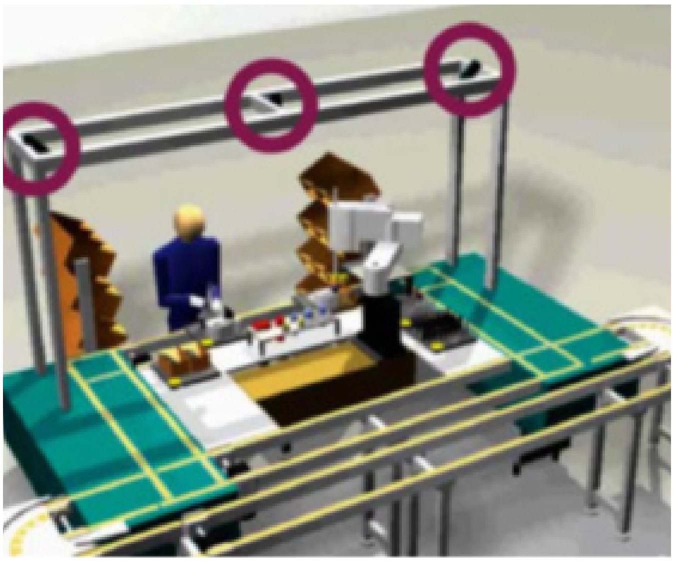

Figure 5: The location of the surveillance cameras at the co-operative work place [110].

The use of a single camera in large scale assembly tasks such as the engine assembly is another challenging issue, due to the incapability of one camera to provide all target features of a large workpiece in the field of view (FOV) [116]. In such these tasks, the sensors should have the capability to work in both long and small ranges, and also because of the large size for assembly objects, which make them exposed to the influences of the gravity and inertia, and the structure of the assembly system for the object in comparison with the one for small objects during the repetition of the assembly task [117].

To cope with the challenge of using the sensory techniques in large-scale assembly tasks, proposed using a multisensory system (two cameras and two laser distance sensors) for achieving high assembly accuracy (0.2 mm in translation) for a large scale object (710×625 mm) [117].

Another challenge with of vision sensors is in the contact assembly tasks (whether the vision system used is a single camera, multi cameras, or some kinds of 3-D reconstructions) is how to cope with the blurring images which are resulted from the closeness of the camera to the object. Therefore, force sensor based methods, such as introduced by Payeur [118,119], are used to compensate for the shortcomings of the vision sensor as force sensors can work in the stage of contacting the object. This leads to that the force sensor based methods cannot be used for guidance from a remote position. therefore, in order to combine the advantages of these types of sensors, hybrid sensor-based methods such as vision-force sensor based have been introduced [120].

The cooperation of ENH camera with a laser pointer with a new IBVS approach are used by Xie [121] for detecting, grasping and assembly metallic parts, the importance of their work in in the robustness of their proposed technique to the camera calibration and hand-eye error.

Vision sensor have also been used with assembly robots in on-site services or construction processes, where the robots perform the assembly of prefabricated components with simple geometric shapes, such as bricks and blocks. One of the main goals for the use of robots in these tasks is to perform rapid and efficient planning for various assembly tasks. However, the unification and calibration of Cartesian coordinates in world units, involving the assembly robot, construction site and the assembly process is not an easy job. To cope with this issue, a visual calibration technique, such as the"ENH"or"E2H,"can assist the articulated robot to establish autonomous positioning systems [122]. 3D laser scanning methods can also be used for calculating position by reconstructing the space [123].

Ding used a vision system (namely Sony a5100) with a 6 DOF robot (namely ABB IRB6700-235) for assembling of construction bricks in three different scenarios which are wall, stair and pyramids. The construction area was around 8 m^2. The camera was used to capture 48 images for the whole area, including the robot end-effector and the breaks, these images were used for the scene construction. An image-based model was used for calibrating robot poses in order to unify of the robot coordinate system and the construction area. The results showed how their automatic assembly system is time-efficient compared with the manual method that frequently requires human interventions.

Another challenge with robotic assembly tasks is on how to define the priority for parts or components to be inserted. Any assembly component must be assembled in a certain number of sequences. However, the increment in the number of components will lead to an increment in the number of assembly sequences, and thus makes the assembly process a very time consuming.

In order to cope with this problem, Mishra used a combination of an overhead CCD camera (namely Basler acA1300-22gc GigE camera with a resolution of 1.3 megapixels), NI vision toolbox, a 6 DOF robot (namely Yaskawa Motoman MH5) and a Sexual Genetic Algorithm (SGA) for obtaining an optimum assembly sequence [124]. The researchers preferred using an overhead camera for three reasons: 1) the task requires assembling of two parts positioned in two different locations, 2) the use of E2H setup provides a panoramic view and a wide FOV of the workspace, and 3) the assembly task requires clearing of the camera's FOV by moving the robot away after the completion of its motion sequences, this action could not be done if the camera was fixed in any location except for the selected position.

2. SUMMARY

The use of vision sensors in industrial robotic applications has gained increasing attention from researchers due to their robustness and ever reducing hardware cost compared with other measuring systems such as laser trackers. The literature review shows that vision systems can drive the robot to perform machining tasks with an acceptable accuracy for most general applications. However, there are some factors that limit the use of vision sensors in many industrial robotic applications, for instance, in the pick and place tasks, a single FOV camera cannot be used to view and track a robot moving in a large measurement area. Also, one camera cannot provide sufficient visual information for both the manipulator and the target to be grasped or assembled at the same time. In order to cope with this limitation, a multi cameras system was suggested which may not be a practical solution for tracking of robots due to 1) the requirements of manual or automatic computation of the relationship between the camera views, and 2) the high hardware cost [125]. The problem of the hardware cost could be partially solved by using cheap sensors after enhancing their resolution via Super Resolution techniques Alzarok [126].

The high computational cost of vision sensors is another concern in robotic grasping and assembly tasks. In order to reduce this cost, the researchers have proposed the use of a small observation window instead of utilising the full resolution of the camera. Another solution is to implementing the algorithm on a separate hardware system such as an FPGA card. The problem of the computational cost for vision sensors restricts the performance of the industrial robots in these applications, for instance, the travelling speed of the robot during real-time tracking process must be kept low in order to give sufficient time for capturing and processing the captured frames. This problem has more effect in the robotic grasping tasks which require the ability of cameras to track the motion of the robot and a moving target; and allow the robot to grasp the target at a certain point (the interception time). However, using techniques such as artificial networks might improve the response of the robotic system. Moreover, the enhancements made in the imaging processing algorithms such as using the hybrid Kalman filter method can also enhance the tracking performance for the vision system.

There are also other challenges which partially don't relate to the vision sensors, such as those related to the target object (e.g., size, motion scenario and speed value). The small-sized objects with similar shapes such as screws in micro assembly, those can't be recognised from each other without knowing of their size, the late requires positioning of the camera very close to the objects for narrow FOV. Moreover, objects with big sizes brings gravity and inertia problems for the robotic system. The appearance of the object is another vital matter in the detection process. For instance, the objects a with metallic surface cannot be recognised by cameras due to the difficulty of its segmentation from the background.

The measurement area is another issue, if it has a big volume, then a multi cameras system with an extra computational cost is required to successfully perform the task, and if has a small volume, then high resolution cameras with an extra hardware cost is needed.

The unprepared environment also brings another challenge for the vision systems. Unwanted objects that might exist in the measurement volume could make that visual detection process very hard job, especially if they share similar features (such as colour and shape) with the target object. Moreover, human operators, especially in manufacturing assembly lines (cobots) also could bring the collision problem and badly affect the detection and tracking process. The bright lights in the industrial locations can cause blooming and glare which lead to interference with tracking, colour classification, and other processing.

3. CONCLUSION

During the last two decades vision techniques have been developed into a very useful technology within industrial robotic applications. This paper has reviewed the contribution of vision sensors for improving the performance of industrial robots in machining applications such as the pose estimation for the industrial robots, grasping of objects, and assembly tasks. Each application area has been explored thoroughly with an insight to the related research work conducted. This paper also highlights the advancements in the field of vision sensors and image processing algorithms which can enhance the performance of industrial robots. It is envisaged that the progress in the vision technology will increase the future uses of articulated robots in high-speed machining tasks. In the aforementioned automation tasks, the vision sensor is required to identify and track moving targets to be grasped or assembled, the tracked target can be the object to be manipulated by the robot for example from the conveyor or it mounted on a stationary workpiece. From the reviewed papers, it can be concluded that there are two main performance requirements for applying vision systems in grasping and assembly tasks which are: (1) high accuracy of the output outcomes and (2) a quick response of the visual tracking system. In other words, the accuracy and time efficiency are both important in these applications. However, achieving the two requirements together is still very challenging due to the fact that more precise tracking tasks are often performed at longer processing time, while quicker responses for the tracking system are more exposed to errors. Performance improvements for the visual tracking systems, in terms of achieving the two aforementioned requirements without trade-off between them can be further suggested.

REFERENCES

[1] Standard, I. Manipulating industrial robots–vocabulary (1994).

[2] Aalsalem, M.Y., Khan, W.Z., Arshad, Q.A. A low cost vision based hybrid fiducial mark tracking technique for mobile industrial robots. ArXiv preprint arXiv (2012).

[3] Reinhart, G., Gräser, R.-G., Klingel, R. Qualification of standard industrial robots to cope with sophisticated assembly tasks. CIRP Annals-Manufacturing Technology (1998) 47(1): 1-4.

[4] Young, K., Pickin, C.G. Accuracy assessment of the modern industrial robot. Industrial Robot: An International Journal (2000) 27(6): 427-436.

[5] Pérez, L., et al. Robot guidance using machine vision techniques in industrial environments: a comparative review. Sensors (2016) 16(3): 335.

[6] Yuan, J., Yu, S. End-effector position-orientation measurement. IEEE Transactions on Robotics and Automation (1999) 15(3): 592-595.

[7] Zhang, L., et al. Tracking mobile robot in indoor wireless sensor networks. Mathematical Problems in Engineering (2014).

[8] Corke, P., Lobo, J., Dias, J. An introduction to inertial and visual sensing. The International Journal of Robotics Research (2007) 26(6): 519-535.

[9] Kiang, C.T., Spowage, A., Yoong, C.K. Review of control and sensor system of flexible manipulator. Journal of Intelligent & Robotic Systems (2015) 77(1): 187-213.

[10] Axelsson, P. Sensor fusion and control applied to industrial manipulators (2014).

[11] Hasan, A., et al. Trajectory tracking for a serial robot manipulator passing through singular configurations based on the adaptive kinematics jacobian method. Proceedings of The Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering (2009) 223(3): 393-415.

[12] García-Rodríguez, R., Parra-Vega, V. Cartesian sliding PID control schemes for tracking robots with uncertain Jacobian. Transactions of The Institute of Measurement and Control (2011).

[13] Borenstein, J., Feng, L., Everett, H. Navigating mobile robots: systems and techniques (1996).

[14] Wang, Z., et al. Experimental comparison of dynamic tracking performance of iGPS and laser tracker. The International Journal of Advanced Manufacturing Technology (2011) 56(1-4): 205-213.

[15] Franceschini, F., et al. Distributed large-scale dimensional metrology: new insights. Springer Science & Business Media (2011).

[16] Müller, T., Schwendemann, J. IGPS–ein vielseitiges messsystem hoher genauigkeit. Allg Vermess Nachr (2009)4: 146-157.

[17] Second, A. Constellation 3D-i: error budget and specifications. White Paper 063102. Arc Second. Inc., Dulles, VA (2002).

[18] Lin, C.Y., Wang, C., Tomizuka, M. Pose estimation in industrial machine vision systems under sensing dynamics: A statistical learning approach. In Robotics and Automation (ICRA), 2014 IEEE International Conference on (2014).

[19] Wang, C., Chen, W., Tomizuka, M. Robot end-effector sensing with position sensitive detector and inertial sensors. In Robotics and Automation (ICRA), 2012 IEEE International Conference on (2012).

[20] Nakamura, T. Real-time 3-D object tracking using Kinect sensor. In Robotics and Biomimetics (ROBIO), 2011 IEEE International Conference on (2011).

[21] Jurado, F., Palacios, G., Flores, F. 3D color-based tracking system for real-time using kinect sensor. in SOMIXXVII Congreso de Instrumentación (2012).

[22] Angelidis, A., Vosniakos, G.C. Prediction and compensation of relative position error along industrial robot end-effector paths. International Journal of Precision Engineering and Manufacturing (2014) 15(1): 63-73.

[23] M&M. Machine vision market by deployment (general & robotic cell), component (hardware and software), product (pc-based and smart camera-based), application, end-user industry & region - forecast till 2025 2020, Markets and Markets Resarch Ltd: India.

[24] Norman, A.R., et al. Validation of iGPS as an external measurement system for cooperative robot positioning. The International Journal of Advanced Manufacturing Technology (2013) 64(1-4): 427-446.

[25] Maisano, D.A., et al. Indoor GPS: system functionality and initial performance evaluation. International Journal of Manufacturing Research (2008) 3(3): 335-349.

[26] Schmitt, R., et al. Performance evaluation of iGPS for industrial applications. In Proceedings of the international conference on indoor positioning and indoor navigation–IPIN (2010).

[27] Muelaner, J.E., et al. Verification of the indoor GPS system by comparison with points calibrated using a network of laser tracker measurements. In Proceedings of the 6th CIRP-Sponsored International Conference on Digital Enterprise Technology Springer (2010).

[28] Schmitt, R., Schönberg, A., Damm, B. Indoor-GPS based robots as a key technology for versatile production. In Robotics (ISR), 2010 41st International Symposium on and 2010 6th German Conference on Robotics (ROBOTIK) (2010): VDE.

[29] Nam, S.H., Oh, S.Y. Real-time dynamic visual tracking using PSD sensors and extended trapezoidal motion planning. Applied Intelligence (1999) 10(1): 53-70.

[30] Benedetti, E., et al. Exploiting performance of different low-cost sensors for small amplitude oscillatory motion monitoring: preliminary comparisons in view of possible integration. Journal of Sensors (2016).

[31] Huang, S.J., Tsai, J.P. Robotic automatic assembly system for random operating condition. The International Journal of Advanced Manufacturing Technology (2005) 27(3-4): 334-344.

[32] Cannons, K. A review of visual tracking. Dept. Comput. Sci. Eng., York Univ., Toronto, Canada, Tech. Rep. CSE-2008-07 (2008).

[33] Jalal, A.S., Singh, V. The state-of-the-art in visual object tracking. Informatica (2012) 36(3).

[34] Dubuisson, S., Gonzales, C. A survey of datasets for visual tracking. Machine Vision and Applications (2016) 27(1): 23-52.

[35] Han, Y., Hahn, H. Visual tracking of a moving target using active contour based SSD algorithm. Robotics and Autonomous Systems (2005) 53(3): 265-281.

[36] Bishop, B.E., Spong, M.W. Adaptive calibration and control of 2D monocular visual servo systems. Control Engineering Practice (1999) 7(3): 423-430.

[37] Carelli, R., Nasisi, O., Kuchen, B. Adaptive robot control with visual feedback. In American Control Conference (1994).

[38] Cheah, C.C., et al. Approximate jacobian control for robots with uncertain kinematics and dynamics. Robotics and Automation, IEEE Transactions on (2003) 19(4): 692-702.

[39] Chaumette, F. Potential problems of stability and convergence in image-based and position-based visual servoing, in The confluence of vision and control (1998): 66-78.

[40] Chen, J., et al. Adaptive homography-based visual servo tracking. In Intelligent Robots and Systems, 2003. Proceedings. 2003 IEEE/RSJ International Conference on (2003).

[41] Deng, L., Janabi-Sharifi, F., Wilson, W.J. Stability and robustness of visual servoing methods. In Robotics and Automation, 2002. Proceedings. ICRA'02. IEEE International Conference on (2002).

[42] Espiau, B., Chaumette, F., Rives, P. A new approach to visual servoing in robotics. Robotics and Automation, IEEE Transactions on (1992) 8(3): 313-326.

[43] Hutchinson, S., Hager, G.D., Corke, P.I. A tutorial on visual servo control. Robotics and Automation, IEEE Transactions on (1996) 12(5): 651-670.

[44] Lippiello, V., Siciliano, B., Villani, L. Eye-in-hand/eye-to-hand multi-camera visual servoing. In Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC'05. 44th IEEE Conference on (2005).

[45] Assa, A., Janabi-Sharifi, F. Virtual visual servoing for multicamera pose estimation (2015).

[46] Malis, E., Morel, G., Chaumette, F. Robot control using disparate multiple sensors. The International Journal of Robotics Research (2001) 20(5): 364-377.

[47] Flandin, G., Chaumette, F., Marchand, E. Eye-in-hand/eye-to-hand cooperation for visual servoing. In Robotics and Automation, 2000. Proceedings. ICRA'00. IEEE International Conference on (2000).

[48] Marchand, E., Hager, G.D. Dynamic sensor planning in visual servoing. In Robotics and Automation, 1998. Proceedings. 1998 IEEE International Conference on: IEEE (1998).

[49] Kragic, D., Christensen, H.I. Survey on visual servoing for manipulation. Computational Vision and Active Perception Laboratory, Fiskartorpsv (2002) 15.

[50] Malis, E., Chaumette, F., Boudet, S. 2½D visual servoing. Robotics and Automation, IEEE Transactions on (1999) 15(2): 238-250.

[51] Elena, M., et al. Variable structure PID Controller for cooperative eye-in-hand/eye-to-hand Visual Servoing. In Control Applications, 2003. CCA 2003. Proceedings of 2003 IEEE Conference on (2003).

[52] Hartley, R., Zisserman, A. Multiple view geometry in computer vision. Cambridge University Press (2003).

[53] Yoon, Y., DeSouza, G.N., Kak, A.C. Real-time tracking and pose estimation for industrial objects using geometric features. In Robotics and Automation, 2003. Proceedings. ICRA'03. IEEE International Conference on (2003).

[54] Harris, C. Tracking with rigid models. In Active Vision: MIT Press (1993).

[55] Marchand, E., et al. Robust real-time visual tracking using a 2D-3D model-based approach. in IEEE Int. Conf. on Computer Vision, ICCV'99 (1999).

[56] Jia, B., et al. Visual trajectory tracking of industrial manipulator with

iterative learning control. Industrial Robot: An International Journal (2015) 42(1).

[57] Hua, C., Wang, Y., Guan, X. Visual tracking control for an uncalibrated robot system with unknown camera parameters. Robotics and Computer-Integrated Manufacturing (2014) 30(1): 19-24.

[58] Luo, R.C., Mullen, R., Wessell, D.E. An adaptive robotic tracking system using optical flow. In Robotics and Automation, International Conference on (1988).

[59] Wang, H., Liu, Y. H, Chen, W. Visual tracking of robots in uncalibrated environments. Mechatronics (2012) 22(4): 390-397.

[60] Wang, H., Liu, Y.H., Zhou, D. Dynamic visual tracking for manipulators using an uncalibrated fixed camera. Robotics, IEEE Transactions on (2007) 23(3): 610-617.

[61] Chang, W.C., Chai, M.L., Real-time vision-based contour following with laser pointer. In Robotics and Automation, 2003. Proceedings. ICRA'03. IEEE International Conference on (2003).

[62] Wang, H., Liu, Y.H., Adaptive image-based trajectory tracking of robots. In Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on (2005).

[63] Grosso, E., et al. Robust visual servoing in 3-D reaching tasks. Robotics and Automation, IEEE Transactions on (1996) 12(5): 732-742.

[64] Kragic, D., Christensen, H.I. Integration of visual cues for active tracking of an end-effector. In IROS (1999).

[65] Chien, M.C., Huang, A.C. Fat-based adaptive visual servoing for robots with time varying uncertainties.

[66] Lund, H.H., de Ves Cuenca, E., Hallam, J. A simple real-time mobile robot tracking system. Citeseer (1996).

[67] Gupta, O.K., Jarvis, R.A. Robust pose estimation and tracking system for a mobile robot using a panoramic camera. In Robotics Automation and Mechatronics (RAM), 2010 IEEE Conference on (2010).

[68] Baek, S., et al. A robot endeffector tracking system based on feedforward neural networks. Robotics and Autonomous Systems (1999) 28(1): 43-52.

[69] Wang, H., Liu, Y.H., Zhou, D. Adaptive visual servoing using point and line features with an uncalibrated eye-in-hand camera. Robotics, IEEE Transactions on (2008) 24(4): 843-857.

[70] Chang, W.C. Binocular vision-based 3-D trajectory following for autonomous robotic manipulation. Robotica (2007) 25(05): 615-626.

[71] Chang, W.C., Wu, C.H. Hand-eye coordination for robotic assembly tasks. International Journal of Automation and Smart Technology (2012) 2(4): 301-308.

[72] Allen, P.K., et al. Automated tracking and grasping of a moving object with a robotic hand-eye system. Robotics and Automation, IEEE Transactions on (1993) 9(2): 152-165.

[73] Borg, J., et al. Navigation-guidance-based robotic interception of moving objects in industrial settings. Journal of Intelligent and Robotic Systems (2002) 33(1): 1-23.

[74] Borg, J., et al. An ideal proportional navigation guidance system for moving object interception-robotic experiments. In Systems, Man, and Cybernetics, 2000 IEEE International Conference on (2000).

[75] Buttazzo, G.C., Allotta, B., Fanizza, F.P. Mousebuster: a robot for real-time catching. Control Systems, IEEE (1994) 14(1): 49-56.

[76] Keshmiri, M., Xie, W.F. Catching moving objects using a Navigation Guidance technique in a robotic Visual Servoing system. In American Control Conference (ACC) (2013).

[77] Fuentes-Pacheco, J., Ruiz-Ascencio, J., Rendón-Mancha, J. Binocular visual tracking and grasping of a moving object with a 3D trajectory predictor. Journal of Applied Research and Technology (2009) 7(3): 259-274.

[78] Nomura, H., Naito. T. Integrated visual servoing system to grasp industrial parts moving on conveyer by controlling 6DOF arm. In Systems, Man, and Cybernetics, 2000 IEEE International Conference on (2000).

[79] Gupta, G.S., et al. Identification and prediction of a moving object using real-time global vision sensing. In Instrumentation and Measurement Technology Conference, 2003. IMTC'03. Proceedings of the 20th IEEE (2003).

[80] Gupta, G.S., et al. Defect analysis of grit-blasted or spray-painted surface using vision sensing techniques. In Proceedings Image and Vision Computing New Zealand (2003).

[81] Shaw, J., Cheng, K. Object identification and 3-D position calculation using eye-in-hand single camera for robot gripper. In 2016 IEEE International Conference on Industrial Technology (ICIT) (2016).

[82] Ge, L., Jie, Z. A real-time stereo visual servoing for moving object grasping based parallel algorithms. In Industrial Electronics and Applications, 2007. ICIEA 2007. 2nd IEEE Conference on (2007).

[83] Gengenbach, V., et al. Automatic dismantling integrating optical flow into a machine vision-controlled robot system. In Robotics and Automation, 1996. Proceedings., 1996 IEEE International Conference on (1996).

[84] Kim, K., et al. Vision based bin picking for industrial robot (2014).

[85] Guo, H., et al. Real-time detection and classification of machine parts with embedded system for industrial robot grasping. In 2015 IEEE International Conference on Mechatronics and Automation (ICMA) (2015).

[86] Harada, K., et al. Pick and place planning for dual-arm manipulators. In Robotics and Automation (ICRA), 2012 IEEE International Conference on (2012).

[87] Saxena, A., et al. Learning to grasp novel objects using vision. In 10th International Symposium of Experimental Robotics (ISER) Citeseer (2006).

[88] Distante, C., Anglani, A., Taurisano, F. Target reaching by using visual information and Q-learning controllers. Autonomous Robots (2000) 9(1): 41-50.

[89] Cowan, N.J., Weingarten, J.D., Koditschek, D.E. Visual servoing via navigation functions. IEEE Transactions on Robotics and Automation (2002) 18(4): 521-533.

[90] Rajpar, A.H., et al. Location and tracking of robot end-effector based on stereo vision. In Robotics and Biomimetics, 2006. ROBIO'06. IEEE International Conference on (2006).

[91] Berscheid, L., Meißner, P., Kröger, T. Robot learning of shifting objects for grasping in cluttered environments. arXiv preprint arXiv:1907.11035 (2019).

[92] Choi, C., et al. Learning object grasping for soft robot hands. IEEE Robotics and Automation Letters (2018) 3(3): 2370-2377.

[93] Alzarok, H., Fletcher, S., Longstaff, A.P. 3D visual tracking of an articulated robot in precision automated tasks. Sensors (2017) 17(1): 104.

[94] Jain, R.K., Majumder, S., Dutta, A. SCARA based peg-in-hole assembly using compliant IPMC micro gripper. Robotics and Autonomous Systems (2013) 61(3): 297-311.

[95] Okumura, S., Take, N., Okino, N. Error prevention in robotic assembly tasks by a machine vision and statistical pattern recognition method.

International Journal of Production Research (2005) 43(7): 1397-1410.

[96] Jayaweera, N., Webb, P. Adaptive robotic assembly of compliant aero-structure components. Robotics and Computer-Integrated Manufacturing (2007) 23(2): 180-194.

[97] Ramachandram, D., Rajeswari, M. Neural network-based robot visual positioning for intelligent assembly. Journal of Intelligent Manufacturing (2004) 15(2): 219-231.

[98] Sahu, O.P., et al. Part recognition using vision and ultrasonic sensor for robotic assembly system. in 2015 IEEE Student Conference on Research and Development (2015).

[99] Zhen, Y., et al. An algorithm of monocular vision for spatial position based on the mirror image. In Communication Software and Networks, 2011 IEEE 3rd International Conference on (2011).

[100] Li, H., et al. Micro-table posture measuring based on binocular vision. In Electronic Measurement & Instruments, 2009. ICEMI'09. 9th International Conference on (2009).

[101] Zhang, Z., Zhang, J., Xu, D. Design of microassembly system and research on coarse-to-fine alignment strategy in combination with active zooming. In Robot Vision, 2013 IEEE Workshop on (2013).

[102] Liu, S., et al. High precision automatic assembly based on microscopic vision and force information. IEEE Transactions on Automation Science and Engineering (2016) 13(1): 382-393.

[103] Xu, D., et al. Characteristic of monocular microscope vision and its application on assembly of micro-pipe and micro-sphere. In Control Conference (CCC), 2013 32nd Chinese (2013).

[104] Xing, D., Xu, D., Liu, F. Collision detection for blocking cylindrical objects. In Intelligent Robots and Systems, 2015 IEEE/RSJ International Conference on (2015).

[105] Xing, D., et al. Precision assembly among multiple thin objects with various fit types. IEEE/ASME Transactions on Mechatronics (2016) 21(1): 364-378.

[106] Hsu, H.C., et al. Position Control and novel application of SCARA robot with vision system. Advances in Technology Innovation (2017) 2(2): 40-45.

[107] Ahn, C.K., Lee, M.C. An off-line automatic teaching by vision information for robotic assembly task. In Industrial Electronics Society, 2000. IECON 2000. 26th Annual Confjerence of the IEEE (2000).

[108] Davies, E., et al. Machine vision approach for robotic assembly. Assembly Automation (2005) 25(3): 204-216.

[109] Consiglio, S., Seliger, G., Weinert, N. Development of hybrid assembly workplaces. CIRP Annals-Manufacturing Technology (2007) 56(1): 37-40.

[110] Krüger, J., et al. Image based 3D surveillance for flexible man-robot-cooperation. CIRP Annals-Manufacturing Technology (2005) 54(1): 19-22.

[111] Akella, P., et al. Cobots for the automobile assembly line. In Robotics and Automation, 1999. Proceedings. 1999 IEEE International Conference on (1999).

[112] Hu, S.J., et al. Assembly system design and operations for product variety. CIRP Annals-Manufacturing Technology (2011) 60(2): 715-733.

[113] Lippiello, V., Siciliano, B., Villani, L. Interaction control of robot manipulators using force and vision. International Journal of Optomechatronics (2008) 2(3): 257-274.

[114] De Santis, A., et al. Human-robot interaction control using force and vision. In Advances in Control Theory and Applications (2007) 51-70.

[115] Cherubini, A., et al. Collaborative manufacturing with physical human–robot interaction. Robotics and Computer-Integrated Manufacturing (2016) 40: 1-13.

[116] Zhao, Y. Visual servoing for robotic positioning and tracking systems. Citeseer (2012).

[117] Qin, Z., et al. Precise robotic assembly for large-scale objects based on automatic guidance and alignment. IEEE transactions on instrumentation and measurement (2016) 65(6): 1398-1411.

[118] Skubic, M., Volz, R.A. Acquiring robust, force-based assembly skills from human demonstration. IEEE Transactions on Robotics and Automation (2000) 16(6): 772-781.

[119] Payeur, P., et al. Intelligent haptic sensor system for robotic manipulation. IEEE Transactions on Instrumentation and Measurement (2005) 54(4): 1583-1592.

[120] Chang, W.C., Shao, C.K. Hybrid fuzzy control of an eye-to-hand robotic manipulator for autonomous assembly tasks. In SICE Annual Conference 2010, Proceedings of IEEE (2010).

[121] Xie, W.F., et al. Switching control of image-based visual servoing with laser pointer in robotic manufacturing systems. IEEE Transactions on Industrial Electronics (2009)56(2): 520-529.

[122] Feng, C., et al. Vision guided autonomous robotic assembly and as-built scanning on unstructured construction sites. Automation in Construction (2015) 59: 128-138.

[123] Dörfler, K., et al. Mobile robotic brickwork. In Robotic Fabrication in Architecture, Art and Design (2016) 204-217.

[124] Mishra, A., et al. Development of a flexible assembly system using industrial robot with machine vision guidance and dexterous multi-finger gripper, In Precision Product-Process Design and Optimization (2018) 31-71.

[125] Yilmaz, A., Javed, O., Shah, M. Object tracking: a survey. Acm Computing Surveys (2006) 38(4): 13.

[126] Alzarok, H., et al. Investigation of a new method for improving image resolution for camera tracking applications (2015).