A Comparative Review of Hand-Eye Calibration Techniques for Vision Guided Robots

Enebuse, Ikenna; Foo, Mathias; Ibrahim, Babul Salam Ksm Kader; Ahmed, Hafiz; Supmak, Fhon; Eyobu, Odongo Steven

IEEE ACCESS

DOI:
10.1109/ACCESS.2021.3104514

Published: 12/08/2021

Publisher's PDF, also known as Version of record

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 23, 2021, accepted August 6, 2021, date of publication August 12, 2021, date of current version August 19, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3104514

A Comparative Review of Hand-Eye Calibration Techniques for Vision Guided Robots

IKENNA ENEBUSE[1], MATHIAS FOO[2], (Member, IEEE), BABUL SALAM KSM KADER IBRAHIM[3], (Member, IEEE), HAFIZ AHMED[4], (Senior Member, IEEE), FHON SUPMAK[5], AND ODONGO STEVEN EYOBU[6]

[1]Institute for Future Transport and Cities, Coventry University, Coventry CV1 5FB, U.K.
[2]School of Engineering, University of Warwick, Coventry CV4 7AL, U.K.
[3]School of Mechanical, Aerospace and Automotive Engineering, Coventry University, Coventry CV1 5FB, U.K.
[4]Nuclear Futures Institute, Bangor University, Bangor LL57 2DG, U.K.
[5]Oxford Vision and Sensor Technology, Coventry CV4 7EZ, U.K.
[6]Geo-spatial Data and Computational Intelligence Laboratory, School of Computing and Informatics Technology, Makerere University, Kampala, Uganda

Corresponding author: Odongo Steven Eyobu (odongo.eyobu@mak.ac.ug)

The work of Hafiz Ahmed was supported by the Sêr Cymru programme by Welsh European Funding Office (WEFO) through the European Regional Development Fund (ERDF).

ABSTRACT Hand-eye calibration enables proper perception of the environment in which a vision guided robot operates. Additionally, it enables the mapping of the scene in the robots frame. Proper hand-eye calibration is crucial when sub-millimetre perceptual accuracy is needed. For example, in robot assisted surgery, a poorly calibrated robot would cause damage to surrounding vital tissues and organs, endangering the life of a patient. A lot of research has gone into ways of accurately calibrating the hand-eye system of a robot with different levels of success, challenges, resource requirements and complexities. As such, academics and industrial practitioners are faced with the challenge of choosing which algorithm meets the implementation requirements based on the identified constraints. This review aims to give a general overview of the strengths and weaknesses of different hand-eye calibration algorithms available to academics and industrial practitioners to make an informed design decision, as well as incite possible areas of research based on the identified challenges. We also discuss different calibration targets, which is an important part of the calibration process that is often overlooked in the design process.

INDEX TERMS Calibration target, camera-world transform, computer vision, hand-eye calibration, robot-hand transform, rotation, translation, vision guided robot.

I. INTRODUCTION

Industrial robots have been around for decades, first gaining popularity in the automotive industry [1]. Automotive plants were suitable for early industrial robots because the tasks in these plants involve a high level of repeatability, large payloads, and moderate speeds. Robots are also being used in a growing number of sectors, such as chicken deboning in the food industry [2]–[4], drug manufacturing in the pharmaceutical industry [5], [6], and aircraft engine construction in the aerospace industry [7]–[9]. According to the International Federation of Robotics (IFR), over 1.7 million new industrial robots will be deployed globally in 2021 [10], and vision systems are now becoming a major component of many industrial robots as they improve the capabilities of robots in operation. For example, vision guided robots can allow for variability in the positioning of work object or deviations in programmed pathway without breaking the production flow [11]–[13].

Emerging applications demand that industrial robots not only be faster, but also be able to accurately identify and find parts that are randomly placed on moving conveyors, containers, or on pallets [14]–[16]. Machine vision systems, which have been around for decades, are now being used in conjunction with robotics to aid automation systems in the processing of such components [17], [18].

Vision guided robotics (VGR) are rapidly becoming a key enabler for the automation of a broad range of processes in a wide range of industries. A typical vision guided robot has a camera attached close to the robot hand or gripper with

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

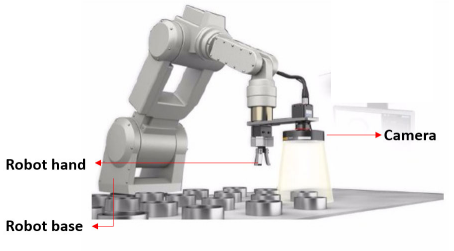


FIGURE 1. Vision guided robot for pick-and-place application [19].

which it can perceive the work environment (Figure 1). The two major areas in the field of computer vision are the 2D and 3D technologies. In a flat plane relative to the robot, a 2D VGR device processes the captured images of randomly place pieces. These images are 2D projections of the 3D spatial pieces, which results in loss of depth information. A 3D VGR device on the other hand, can process parts that are randomly positioned in three dimensions (i.e., X-Y-Z) and can also accurately determine the 3D orientation of each part. In practice, 2D machine vision is typically accomplished using a digital camera and a software that analyses a digital image of the part's 2D position and orientation in preparation for robotic handling or processing [20], [21]. The 3D vision system on the other hand, uses sensors like laser displacement, structured light and stereo camera capable of generating a point cloud representation of a surface in the 3D space [22]–[24]. The point cloud enables the spatial reconstruction of a 3D scene, which facilitates the handling of a wide variety of complex objects in a challenging environment, thereby enhancing the capabilities of robots for vision guided applications. One particular advantage is being able to pickup objects placed on a surface with irregular height, which would be difficult for the 2D vision system.

Applications of vision guided robots include part assembly [25], bin picking [26], inspection [27] etc. These robots can either have the camera mounted in a fixed position with a fixed field of view (eye-to-hand configuration), or have the camera mounted on the hand of the robot (eye-in-hand configuration), so that new images can be acquired by changing the point of view of the camera. However, the robot can only perceive the 3D world based on its own base frame. In order for the robot to obtain an accurate estimate of the 3D position and orientation of a part relative to its own base within the work volume, it is necessary to know the relative position and orientation between the hand and the robot base, between the camera and the hand, and between the object and the camera. These three tasks require the calibration of robot [28]–[30], camera [31], [32], and robot hand-to-camera (hand-eye) [33], [34]. Robot calibration is needed because, despite the fact that robots have very good repeatability, they are poor when it comes to absolute accuracy, due to inherent differences between the ideal and actual kinematic parameters. Camera intrinsic calibration is required to ensure that the images captured are of accurate dimensions and free of lens distortion, which would otherwise introduce errors in the measurement estimates being fed back to the robot during operation. Hand-eye calibration ensures that the measurements made by the camera is converted to the reference used by the robot for measurement. The focus of this review is on hand-eye calibration and its associated challenges to robotic vision system.

The rest of the paper is organised as follows. Section II introduces the problem of hand-eye calibration, Section III discusses the different hand-eye calibration algorithms. A comparative analysis of calibration target is given in Section IV, while common challenges associated with hand-eye calibration is presented in Section V. Finally a conclusion is given in Section VI.

II. HAND-EYE CALIBRATION

The perception of the environment by robot can be accomplished using a camera. This enables navigation and manipulation of objects in an unknown and dynamic environment. This vision system involves the perspective projection and mapping of a 3D world coordinate point onto a 2D image plane, which can be achieved using a pinhole camera model [35] as shown in Figure 2.

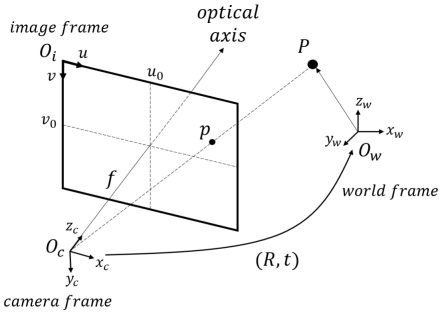


FIGURE 2. Pinhole camera model.

From Figure 2, an optical ray passing through a 3D world point P through the optical centre O_c intersects the image plane at a point p located a distance of f (focal length) from the optical centre. To obtain the point p in the image plane $O_i(u, v)$, the world coordinate points O_w first has to be transformed to the camera coordinate at O_c. This is achieved using the transformation Equation (1). From Equation (1), the camera coordinate points $P^c = (x_c, y_c, z_c)$ are realised from world coordinate points $P^w = (x_w, y_w, z_w)$ using the rigid body homogeneous transformation matrix H_w^c

$$\begin{pmatrix} P^c \\ 1 \end{pmatrix} = H_w^c \begin{pmatrix} P^w \\ 1 \end{pmatrix}, \tag{1}$$

Equation (1) can also be expressed as

$$\begin{pmatrix} x_c \\ y_c \\ z_c \\ 1 \end{pmatrix} = \begin{pmatrix} R_w^c & t_w^c \\ 0_{1\times 3}^T & 1 \end{pmatrix} \begin{pmatrix} x_w \\ y_w \\ z_w \\ 1 \end{pmatrix}, \tag{2}$$

where R_w^c and t_w^c denote rotation and translation, respectively, from the world to camera coordinate frames. These parameters are regarded as the extrinsic parameters of the camera.

The projection of the points in the camera coordinate onto the image plane based on the pinhole camera model is given by Equation (3).

$$\begin{pmatrix} u \\ v \end{pmatrix} = \frac{f}{z_c} \begin{pmatrix} x_c \\ y_c \end{pmatrix}, \tag{3}$$

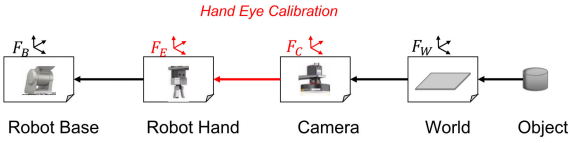


FIGURE 3. Relationships between component frames for vision guided robot.

The task of computing the relative 3D position and orientation between the camera and the robot hand in an eye-on-hand configuration, where the camera is rigidly attached to the robot hand, is known as hand-eye calibration. More specifically, this is the task of computing the relative rotation and translation (homogeneous transformation) between two coordinate frames, one centred at the camera lens centre, and the other at the robot hand. Figure 3 shows the relationships between the different components frame of a vision guided robot operation. To ensure easy operation of the robot, all commands to the robot are referenced to the robot base frame. Hence, for a complete identification of the object based on the robot base frame, all the relationships must be obtained. While the relationship between the robot base and the robot hand can be realised from the robot kinematic model, the relationship between the camera and the world (see Equation (1)) can be obtained from camera calibration. This results in the relationship between the camera and the robot hand need to be computed. This relationship cannot be measured directly because [33]

1) the measurement path may be obstructed by the geometry of the sensor, the robot or other parts of the system
2) the hand and camera frames are unreachable. The camera frame is unreachable because it is the intersection of various link axes while the camera frame is unreachable because its origin is at the focus point inside the camera.

Since the direct measurements are difficult, other approaches have been investigated to solve the problem. Earlier approaches used non-linear optimisation of a model that coupled the robot forward kinematics with the hand-eye system [36]. These techniques are quite expensive computationally and require estimation of a large number of variables. In view of that, the most common technique used [33], [37]–[39] is based on solving the homogeneous transform equation according to Equation (4), where $A_{c_2}^{c_1}$ and $B_{h_2}^{h_1}$ are the homogeneous transform matrices for the motion of the camera and robot hand, respectively, between two positions 1 and 2, and X_h^c is the required robot hand to camera homogeneous transform. If the position and orientation of the hand are known, the position and orientation of the camera can be simply computed, vice versa. The object can then be located with respect to the robot base and locating information from different views can be fused. The first challenge encountered during hand-eye calibration is usually the estimation of the pose of the camera relative to the world as the hand pose can easily be acquired from the robot forward kinematic chain. Depending on how the camera pose is estimated, the hand-eye calibration can be regarded as either target-based or targetless. Target-based hand-eye calibration takes advantage of specially made visual features of known dimensions called calibration objects or target - whose origin is set as the origin of the world frame - to estimate the pose of the camera using special algorithms like the Perspective-n-Point [50]. Targetless hand-eye calibration without a calibration target uses techniques such as in structure from motion [42], [48], tool motion tracking [49] etc, to estimate the pose of the camera with respect to the world. These methods can prove useful when taking the size and weight of the calibration object into considerations as well as the size of the work space for the robot motion. These considerations for a calibration object usually come into play when there is strict limitation of payload of a mobile robot such as in space application, or sterility of the setup in medical applications. In this review, only techniques based on target-based hand-eye calibration are considered.

Furthermore, it is important to note that the methods presented in this review focus primarily on the deterministic formulation. Therefore, this review is by no means an exhaustive list of the approaches to hand-eye calibration for visual guided robots. We note that there are other key methods available, which include (but not limited to) model based [40], [41] and probabilistic [43]–[47] formulation of the hand-eye calibration problem. The intent of this review is to act as a guide to academics and industrial practitioners from which further research in this topic area can be incited.

III. HAND-EYE CALIBRATION ALGORITHMS

A. HOMOGENEOUS TRANSFORM EQUATION

Based on the work of Shiu and Ahmad [33], the hand-eye transform can be obtained by solving the homogeneous transform equation given by

$$A_{c_2}^{c_1} X_h^c = X_h^c B_{h_2}^{h_1}, \tag{4}$$

where, $A_{c_2}^{c_1}$ and $B_{h_2}^{h_1}$ are the homogeneous transform matrices representation of the relative motions of the attached camera and the robot hand between two points, respectively, while X_h^c is the required transform between the robot hand and the camera as shown in Figure 4. $A_{c_2}^{c_1}$ and $B_{h_2}^{h_1}$ can be expressed as the product of two rigid body transform given by

$$A_{c_2}^{c_1} = A_w^{c_1} (A_w^{c_2})^{-1}, \tag{5a}$$

$$B_{h_2}^{h_1} = B_b^{h_1} (B_b^{h_2})^{-1}, \tag{5b}$$

where $A_w^{c_1}, A_w^{c_2}$ and $B_b^{h_1}, B_b^{h_2}$ are the poses of the camera with respect to the world frame or calibration object, and the poses of the robot hand and with respect to the robot base,

respectively, for different robot positions. Equation (4) can be represented in a matrix form as

$$\begin{pmatrix} R_A & \vec{t}_A \\ 0^T & 1 \end{pmatrix} \begin{pmatrix} R_X & \vec{t}_X \\ 0^T & 1 \end{pmatrix} = \begin{pmatrix} R_X & \vec{t}_X \\ 0^T & 1 \end{pmatrix} \begin{pmatrix} R_B & \vec{t}_B \\ 0^T & 1 \end{pmatrix}, \quad (6)$$

where R is a 3×3 rotation matrix and $\vec{t}$ is a 3×1 translation vector. Hence, the calibration operation involves obtaining sets of robot hand and camera poses as shown in Figure 5. While the hand poses can easily be obtained from the robot forward kinematics using the joint encoder readings, the camera pose is usually estimated by observing a set of 3D points provided by a calibration object and their corresponding 2D images using Perspective-n-point algorithm [50], [51]. While this formulation shows a more intuitive way to represent and solve the hand-eye problem, estimating the hand-eye transform based on Equation (4) is not trivial. This is because the Special Euclidean SE(3) group structure of the homogeneous matrices must be preserved in the solution. Hence, the solution to this form of matrix equation using general matrix algebra [52] would not work.

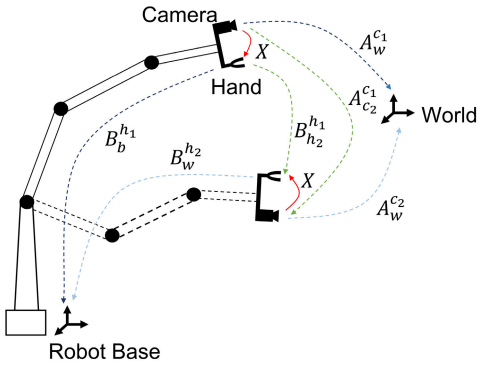


FIGURE 4. Hand-eye calibration setup.

Finding methods of solving the homogeneous transform equation that meet this requirement has been the focus of majority of the research in hand-eye calibration. Several solutions have been proffered over the years, each with its strengths and weaknesses. They can be grouped based on how the rotation and translation parameters are estimated as *separated* or *simultaneous* solutions. In the separated solutions, the rotation parameter is first estimated based on representing Equation (6) as

$$\begin{pmatrix} R_A R_X & R_A \vec{t}_X + \vec{t}_A \\ 0^T & 1 \end{pmatrix} = \begin{pmatrix} R_X R_B & R_X \vec{t}_B + \vec{t}_X \\ 0^T & 1 \end{pmatrix}. \quad (7)$$

hence,

$$R_A R_X = R_X R_B, \quad (8a)$$

$$R_A \vec{t}_X + \vec{t}_A = R_X \vec{t}_B + \vec{t}_X. \quad (8b)$$

If R_X is known, then Equation (8b) becomes linear and $\vec{t}_X$ can then be estimated. The different techniques that focus on the parametrisation of R_X include, Angle-axis [33], [37], Lie algebra [38], Quaternions [39] and Kronecker product [53].

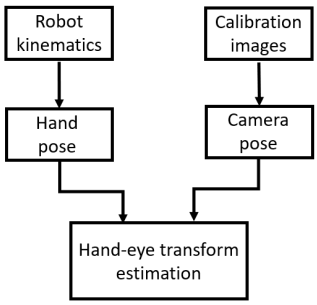


FIGURE 5. Calibration process flow using homogeneous transform equation.

For the details of the implementations of these algorithms, see the listed references above. Based on the practical considerations, generally, this group of solutions is computationally fast but suffers in terms of accuracy, especially in the translation estimates. This is due to the assumption that no relationship exist between the rotation and translation parameters, hence, their separate estimation. However, these two parameters are tightly coupled with high level of nonlinearity [54] and estimating them separately would lead to propagation of errors from the rotation estimates onto the translation estimates.

The simultaneous solutions provide a way of solving for the rotation and translation parameters simultaneously, either analytically or by way of numerical optimisation. Representative implementations based on analytical approach include Quaternions [55], Screw motion [56], Dual Quaternions [57], Kronecker product [58], Dual Tensor [59], and Dual Lie algebra [60], while implementations based on numerical optimisation include Gradient/Newton optimisation method [61], Linear-matrix-inequality [62], Alternative linear programming [63], and Pseudo-inverse [64]. These methods can generate highly accurate results and generally avoid the problem stated earlier for the separated solutions. However, their implementations are usually complex, which may affect their computational speed. Furthermore, the optimisation methods may suffer from the problem of not guaranteeing convergence, being trapped in a local minima of the cost function or dependent on a good starting estimate. A comparison of these approaches based on the accuracy and the computational speed is shown in Table 1. The accuracy criteria is based on the Euclidean norm of the combined rotation and translation error (unitless) for N robot movements derived from Equation (4), as given by Equation (9). The computation time is in seconds, based on execution on a MacBook Pro 2017 with i7-3.5Ghz CPU along with the MATLAB r2018a software [65].

$$\text{Error} = \frac{1}{N} \sqrt{\sum_{i=1}^{N} \|A_i X - X B_i\|^2}. \quad (9)$$

It is important to note that the values in Table 1 can only be considered as an overview of what can be expected, especially due to the fact that it is devoid of any measurement

uncertainty. This is a largely ignored area of research when considering hand-eye calibration, as only a hand-full of works [45], [54], [65], [66] has taken measurement uncertainty into consideration given the large number of research outputs in this area. However, it is important to note that the accuracy of the calibration methods can be improved by increasing the number of robot movements used during the calibration process, maximising the angular spread between the different robot movements, minimising the distance between the camera and the calibration target and minimising the distance moved by the robot arm between two positions [37].

B. REPROJECTION ERROR MINIMISATION

The homogeneous transform equation relies on the hand and camera pose information to estimate the hand-eye transformation. As such, errors in these pose estimates will affect the end result of the calibration. While the hand pose errors can be minimised by calibrating the robot [67], through reprojecting the image of the calibration pattern at each hand position and minimising the error between the real image and the reprojected image, the required hand-eye transform can then be estimated as shown in the process flow in Figure 6.

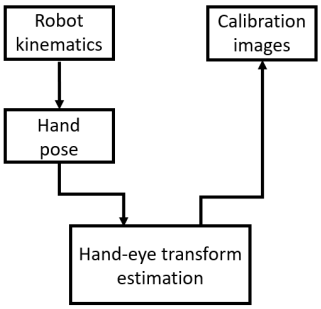


FIGURE 6. Calibration process flow using reprojection error minimisation.

Reprojection error minimisation is a well-known technique used in computer vision for pose estimation [68], [69], 3D measurements [70] and shape reconstruction [71], [72], with high level of accuracy and robustness. It shows how precise an estimated 3D world point $\hat{X}$ recreates the true projection x on the image (see Figure 7). If P is the projection matrix of the camera, then the image projection $\hat{x}$ can be expressed as $\hat{x} = P\hat{X}$, where $e(x, \hat{x})$ represents the reprojection error is the Euclidean distance between x and $\hat{x}$. By minimising e the true projection matrix can be obtained, and if the camera calibration is known, then the pose of the camera can be realised implicitly.

The main advantage of this technique over the homogeneous transform equation is that it directly takes images of the calibration object without requiring explicit pose estimate of the camera, which may otherwise contribute to errors. The Perspective-n-point algorithm is usually used in the estimation of the camera pose information from the pattern images [73], [74]. However, this can be problematic when using cameras with narrow field of view such as in

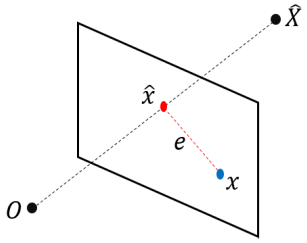


FIGURE 7. Reprojection error.

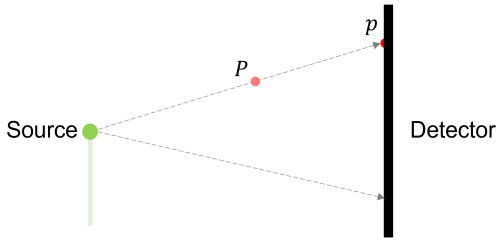


FIGURE 8. Source-detector projection model. Light ray from the source passes through a world object P and is projected on the detector at point p.

thermographic cameras [75]. Furthermore, the formulation of the homogeneous transform equation is perfectly suited to normal cameras, whose optics are modelled using the pinhole camera projection model. When considering vision sensors with different optics, such as in X-rays with source-detector projection model (see Figure 8), it becomes difficult to use the homogeneous transform formulation as the typical pinhole projection model does not provide a proper representation of its optics. One way of achieving this is by using pose graph optimisation [76], which estimates relatives pose of an object based on a network of observed pose sequences. With pose graph optimisation, it becomes possible to extend the calibration to vision sensors with different optical projection model like in source-detector model, where the source pose and the detector pose can be reliably represented in the pose graph.

C. ARTIFICIAL NEURAL NETWORK

Artificial neural network (ANN) is motivated by the neural system in the brain and is one of the most commonly used tools in machine learning [77]. In its basic form, it consists of layers of interconnected nodes, each representing a mathematical function. The strength of ANN comes from its ability to model highly non-linear functions that map an input to an output (see Figure 9). Hence, its application in pattern recognition [78], robotics [79], signal and image processing [80] and non-linear system state estimation [81]–[83] have been very successful. An ANN model is obtained by training the network with a set of input and corresponding output data to obtain a set of optimised network parameters. The trained model with its optimised parameters can then be applied to an appropriate input to get the expected output.

Employing ANN in hand-eye calibration can be thought of as finding a mapping between the hand coordinate with

TABLE 1. Speed and accuracy comparison of different approach to solving the homogeneous transform equation. Speed is in milliseconds while the accuracy is unitless, based on the Euclidean norm of the combined rotation and translation given by equation (9).

Methods	Type	Computation Speed (milliseconds)	Accuracy ($\times 10^{-3}$)
Angle-axis [37]	Separated	40.50	12.046
Lie algebra [38]	Separated	35.95	6.825
Quaternions [39]	Separated	10.19	6.826
Dual quaternions [57]	Simultaneous, Analytical	38.30	6.708
Kronecker product [58]	Simultaneous, Analytical	57.29	7.265
Linear-matrix-inequality [62]	Simultaneous, Optimisation	14.58	7.980
Alternative linear programming [63]	Simultaneous, Optimisation	80.86	7.149
Pseudo-inverse [64]	Simultaneous, Optimisation	8.82	7.613

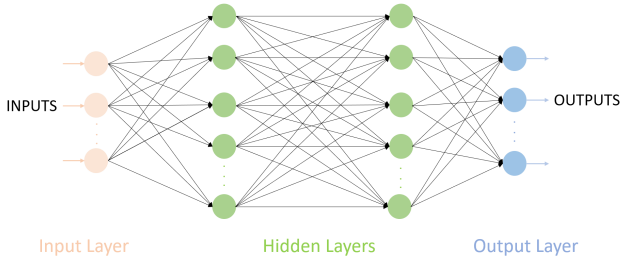


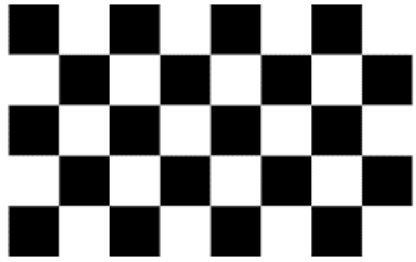
FIGURE 9. A simple artificial neural network with two hidden layers.

respect to the robot base and the respective image coordinate of the calibration object. This problem can be posed as $A = f_n(B)$, where A and B are the robot's hand coordinate and calibration image coordinate respectively and f_n is the function depicting the non-linear ANN model. With a trained model, the required hand coordinate for a corresponding object position as observed from the camera can be obtained. An advantage of this formulation is that it can be used without the knowledge of the camera parameters or pose estimation [84]. This comes from the strong ability of ANN to generalise non-linear relationships between variables, which also makes it suitable for handling noise [84].

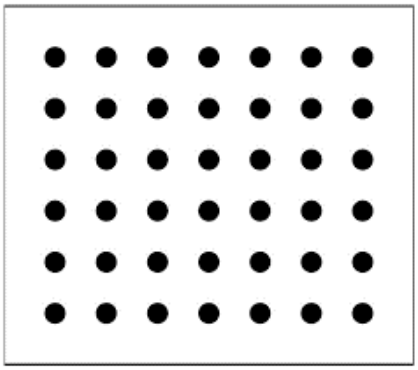
While ANN has some comparative advantages over the methods of homogeneous transform equation and reprojection error minimisation, it is important to note that the solutions provided by ANN are usually unexplainable [85]. This can lead to mistrust of the system and difficulty in troubleshooting problems. Furthermore, the performance of an ANN model is highly dependent on the network structure used [87] for which there is no definitive rule for appropriate specification. As such, it is common to select a network structure based on trial-and-error and users experience.

Parameter over-fitting is another limitation of ANN [88]. This is usually attributed to the failure to properly generalise the model on the available data set, where the model is too simple that could not learn enough, or the model is too complex that it learn too much and over-fits the data. Techniques for preventing over-fitting includes simplifying a complex model, stopping the training early when error starts to increase, data augmentation and regularisation [89].

Table 2 shows a comparison of the methods of homogeneous transform equation, reprojection error minimisation and artificial neural network for hand-eye calibration.



(a) Checkerboard



(b) Circle

FIGURE 10. Different types of calibration patterns [92].

IV. CALIBRATION TARGET

The calibration target (object) is a very important piece of a calibration process, be it for camera calibration or hand-eye calibration. This subject is rarely discussed when dealing with calibration and often times, the decision to use a particular calibration target is not objective with more focus on lens distortion modelling and parameter optimisation [90]. Calibration patterns such as checkerboards and circles are the most used [91] (see Figure 10). This is due to the ease of which they can be created with sufficient accuracy, and their data points can be obtained easily using standard image processing algorithms [91]. During calibration, the calibration pattern captured by the camera can undergo perspective or non-linear distortion, or both. While the perspective distortion is due to the relative 3D position of the points, the non-linear distortion is due to camera lens distortion. How the resulting distortion affects the different calibration patterns determines their reliability.

TABLE 2. Homogeneous transform equation vs reprojection error minimisation vs artificial neural network for hand-eye calibration.

Homogeneous transform equation	Reprojection error minimisation	Artificial neural network
Requires explicit camera pose estimation	Camera pose estimation is implicit	Does not require camera calibration or pose estimation
Suited to camera that can be described with pinhole camera model	Because it uses direct images from the camera, it can accommodate other camera models	Model generalisation means it can accommodate other camera models
No issue with overfitting of solutions	Solutions can be prone to overfitting	Parameter over-fitting can be a limitation
Computation time of 0.142 s [76]	Computation time of 0.272 s [76]	Computation time of 2.355 s [84]
Accuracy of 1.715 mm [76]	Accuracy of 1.380 mm [76]	Accuracy of 0.923 mm [84]

A. CHECKERBOARD TARGET

Checkerboard target is the most common calibration pattern [93]–[96]. The interest points are the corners of the squares, which can be detected as the intersection of the lines that makeup the square edges. Mathematically, these intersection points are the saddle points, which can easily be detected as points, where the first derivative goes to zero. The detection algorithm usually starts by binarising the image, followed by filtering to ensure that the size and organisation meet the dimension and structure specified by the user. The main disadvantage of checkerboard targets is that it is usually difficult to get the exact boundary of the corners [97]. However, the detection of the corners of the squares can usually be done with sufficient level of accuracy because the corners, being infinitely small are mostly invariant to perspective and lens distortion [91].

Because of the alternating colours of adjacent squares, the checkerboard target can be made rotation-invariant by making the number of rows and columns even and odd, respectively, or vise-versa. Otherwise, with both rows and column either even or odd, the pattern creates a 180-degree ambiguity that can be problematic for multi-camera calibration, where similar point needs to be identified by multiple cameras like in the calibration with stereo cameras.

B. CIRCULAR GRID TARGET

Circular grid targets are based on circles with the feature point being at the centre of the circle. Appropriate circles in the target can be detected using characteristics like circularity and convexity, and bad featured circles can be eliminated. While the circles themselves are easy to detect and to be filtered, unlike the checkerboard target, they are not invariant under perspective and lens distortion as shown in Figure 11. Under perspective projection, the circles are imaged as ellipses. Ideally, this can be solved using image rectification, however, the additional lens distortion on the ellipses adds some bias to the detected points, which in general would require a more complex algorithm to correct.

Just as in checkerboard targets, the circular grid target can be made rotation invariant for multi-camera view. This is done by using asymmetric grid pattern as shown in Figure 12.

C. DISCUSSIONS

Checkerboard and circular grid are the most widely used patterns for vision system calibration. The choice of pattern used depends on the application constraints such as accuracy, complexity of detection algorithm, distortion, etc. The feature points for checkerboard pattern are the intersection of the lines that can easily be obtained using standard corner detection strategies [99]. For circular grid patterns, the feature point commonly used is the centre of mass of the circle pixels. Often times the estimated position of checkerboard corner, or circle centre do not fall at the exact point and further computation would be required for sub-pixel accuracy [91]. Generally, the mathematics involved for realising sub-pixel accuracy for circular grids is much more complex than for checkerboard pattern [91]. This complexity is compounded by the fact that the feature point for circular grid is affected by both radial and perspective bias. Hence the accuracy of circular grid depends on how well the true centre of the circle can be determined. Figure 13 illustrates the effect of radial distortion on the accurate detection of the features for checkerboard and circular targets. In this illustration the radial distortion coefficient k_1 as given in Equation (10) is increased from -2 to 2, where $k_i, i = 2, 3, \ldots$ are the distortion coefficients, r_d is the distortion radius, (x_u, y_u) and (x_d, y_d) are the undistorted and distorted image points respectively. Table 3 shows a summary of the comparison between the properties of checkerboard and circular grid patterns.

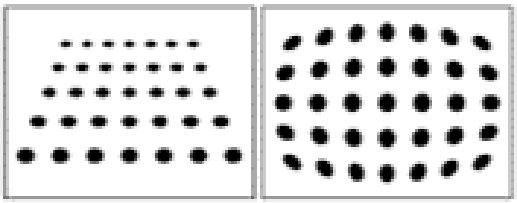


FIGURE 11. Circular grid under perspective (left) and lens (right) distortion [92].

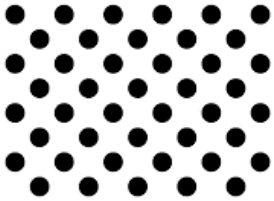


FIGURE 12. Asymmetric circular grid pattern [98].

$$x_u = x_d(1 + k_1 r_d^2 + k_2 r_d^4 + \cdots) \tag{10a}$$

$$y_u = y_d(1 + k_1 r_d^2 + k_2 r_d^4 + \cdots). \tag{10b}$$

TABLE 3. Comparison of checkerboard and circular grid calibration patterns.

	Checkerboard	Circle
Feature points	Corners and edges	Centroid and conics
Detection Complexity	Features can easily be resolved even in the presence of distortion	Feature detection is usually complicated by distortion
Perspective distortion	No	Yes
Effect of lens distortion (See Figure 13)	low	High
Occlusion	It may be possible to interpolate lines to infer intersection	Connecting distorted points to infer intersection is not usually possible
Sub-pixel accuracy	Simple interpolations from already acquired points can be used to achieve sub-pixel accuracy for other points	Special techniques are required to achieve sub-pixel accuracy

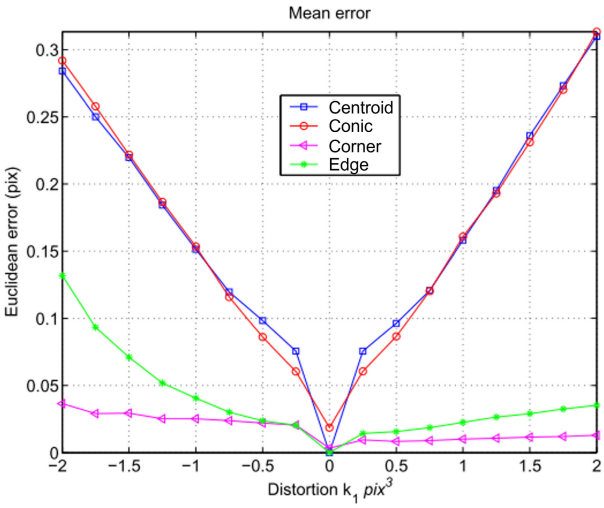


FIGURE 13. Accuracy evaluation for feature points detection for checkerboard (corners and edges) and circular grid (centroid and conics) targets for increasing levels of radial distortion k_1 [91].

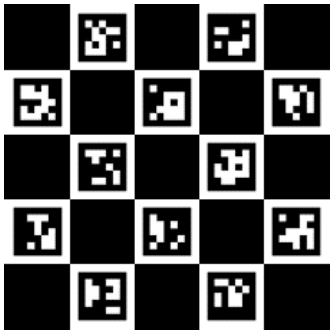


FIGURE 14. CharuCo target [100].

D. OTHER CALIBRATION TARGETS

Other calibration targets exist with the aim of overcoming the limitations of the checkerboard and circular calibration targets. Most of these come with some form of encoding marker. An example of such is the CharuCo target shown in Figure 14.

In the CharuCo target, the light squares are uniquely encoded. Thus, this makes CharuCo target possible to carry out calibration even with part occlusion or poor image conditions such as in inhomogeneous lighting, while maintaining the advantage that the intersection of the square edges or interest points can easily be recovered, when an ordinary checkerboard or circular grid target under these conditions would normally fail. The main drawback of this calibration target is the complex algorithm required for the detection and decoding of the patterns.

V. COMMON CHALLENGES OF HAND-EYE CALIBRATION

Hand-eye calibration is an active field of research in robotics and computer vision mainly due to the importance of precision and accuracy in these industries. For example, while an accuracy level of 1 mm may be required for spot welding operation in the automotive industry, an accuracy measure of at least ten to twenty-fold would be required in the aerospace industry [101]. Similar accuracy levels can also be found in robotic applications in the health industry, where safety is of utmost importance [102]. Achieving this level of accuracy is a major challenge in hand-eye calibration for robots due to a number of factors such as data asynchronicity, noise and limited motion range.

A. DATA ASYNCHRONICITY

The hand-eye calibration problem is constrained on data from two sources: the eye (camera) and the hand (robot). This constraint requires correspondence in the data stream from both sources, which may not be practically possible, resulting in temporal misalignment in the data [103]. This temporal misalignment may be due to the differences in the operating frequency of the sensors, difficulty in synchronising the trigger for the data capture on both sources or missed data in either stream. Many solutions to hand-eye calibration are offline in nature [38], [39], [53], [64], where the calibration setup is made, complete pose data set for both the hand and the eye with respect to the robot base and world, respectively, are acquired, and computation of the required hand-eye transform made. Regardless of the fact that the acquisition of both sets of data are made in discrete steps, data asynchronicity still forms a major problem which affects the correspondence of the data. Offline calibration nonetheless is not suitable for certain applications. An example is in critical operations like robot assisted surgery (RAS), where frequent changes in setup and recalibration is an expensive operation that must be dealt with on the fly [104]. This type of application

require online calibration [105], where data is continually being captured and used to update the calibration algorithm, rendering the need for data synchronisation on both sources very apparent.

One solution to the problem of temporal misalignment is the use of timestamp [106]. By timestamping the data from both sources, users could manually or programmatically synchronise the data streams and also avoid missing data. Cross-correlation techniques can also been used to achieve data synchronisation for hand-eye calibration as in [107], [108]. Normalising and resampling the data before the cross-correlation operation can be used to ensure that differential data length caused by time delay or different sampling rate do not affect the result. A more elegant solution can be found in the use of a real-time embedded operating system for the control of the data capture and synchronisation operation [109]. This, however, would require compatibility with different sensors and robot systems and can quickly make the setup less attractive in terms of cost and complexity.

B. NOISE

Noise is a major problem in hand-eye calibration, which arises as a result of perturbations in the robot-camera assembly. This causes some degree of uncertainty in the calibration results. A direct impact of the noise in hand-eye calibration is the need to use measurements from multiple coordinate frames (greater than the theoretical minimum) for the estimation of the hand-eye transform [37]. The required hand-eye transform is estimated from a system of equations based on the rigid body transform of the robot-camera assembly, which normally results in an overdetermined system [33]. In an ideal scenario, with no noise in the system, because the measurements are physically constrained to be consistent with the robot-camera assembly, the set of equations could be solved by a simple least square method. Since there are more equations than unknowns in the presence of measurement noise, the equations becomes inconsistent and multiple frames or robot motions would be required to accurately estimate the system variables.

Noise in robot hand-eye calibration can be categorised into two forms. These are noise as a result of the robot motion and the camera motions. Noise from the robot motion directly affects the kinematic model of the robot as they are caused by measurements from the joint encoders or optical trackers in the robot. The error in measurement can be due to various factors such as kinematic errors, non-kinematic errors and joint errors.

- *Kinematic Errors:* Kinematic errors are related to and have direct impact on the kinematic model of the robot [110], [111]. These may be due to manufacturing and assembly tolerances, geometry of the robot components such as orthogonality or parallelism or the position of the reference frame.
- *Non-Kinematic Errors:* Unlike kinematic errors, non-kinematic errors are due to the mechanical characteristics of the robot components such as stiffness, blacklash elasticity and impact of temperature [112], [113].
- *Joint Errors:* Joint errors are directly related to the error in motion measured at individual joints of the robots by the joint encoders and are caused by the sensors themselves [114], [115].

Noise from the camera motions is a direct consequence of camera calibration, which can result from low camera quality, poor calibration parameter estimates, low quality calibration pattern etc. These errors, while they can be small from a camera calibration perspective [116], can be propagated to the estimates of the hand-eye calibration.

C. LIMITED MOTION RANGE

The range of allowable motion of the robot hand during calibration have a direct impact on the results of the calibration. Large motions in the robot hand has the effect of suppressing noise in the setup that can arise due to perturbations [37]. Despite this advantage, not every application is able to permit a wide motion range. In RAS, only a small motion range of the surgical tool is permitted. This is usually constrained to within the vicinity of the trocar entry ports [112]. This is done to minimise the damage that can be done to surrounding tissues at the entry ports [117]. In pick-and-place applications like in sorting and assembly facilities, the constraint is the field of view of the camera outside in which the operation of the robot is not feasible [118]. In other instances, the robot motion is limited to a particular area to provide a safe environment in which human operators can operate [119], [120]. In these applications, the robot is controlled by an embedded control system that specifies and limits the motion of the hand to a given work space. While the allowable range of motion of the robot hand cannot always be controlled, a lot of gain can be achieved by implementing proper path planning algorithm and pose selection methods to obtain a well-conditioned robot hand-eye constraints [37], [121], [122].

VI. CONCLUSION

In this review, different solutions to hand-eye calibration were discussed with the aim of presenting their strengths and weaknesses. The purpose was to provide necessary information that would be required for implementation by academics and industrial practitioners, as well as encourage further research. The most common formulation of the problem requires finding a solution to the homogeneous transform equation $AX = XB$. A lot of research have be done in this area, with solutions found using angle-axis representation of the rotation parameter, Lie algebra, Quaternions, Dual Quaternions, Screw motion, Kronecker product as well as using optimisation techniques. Each of the resulting algorithms differ in their level of accuracy and computational requirement, which needs to be taken into account by academics or industrial practitioners depending on their design constraints. Alternate methods that solve the problem using reprojection error minimisation and artificial neural network

are also presented. The main advantages with the method of reprojection error minimisation are error avoidance due to camera pose estimation and the ability to work with camera models (e.g. like source-detection model used for X-rays) other than the pinhole projection model. While the method of artificial neural network also simplifies the problem by using only images taken by the camera, network parametrisation and over-fitting may limit its usage.

Different considerations for choice of calibration pattern are also discussed, with checkerboard and circular grid patterns being the most common calibration targets. While sub-pixel accuracy can be achieved using either of the target choices, circular grid targets usually require more complex algorithms. How the patterns respond to perspective and lens distortion play a huge role their reliability, with circular grid targets being more susceptible to distortions that must be corrected. The CharuCo target on the other hand embeds encodings on its pattern to avoid the limitations of the most commonly used checkered board pattern.

Finally, we discuss about some common challenges that are expected in the calibration of robots eye-hand system. While proper planning and appropriate calibration setup can improve the calibration estimate, it is sometimes difficult to meet all the conditions for improved accuracy and a compromise has to be made. Data asynchronicity, noise and limited motion range are identified as some of the challenges of hand eye calibration that can also gain improvement from proper path planning, calibration setup and robot calibration prior to hand-eye calibration. In general, for accurate vision guided robotic operation, there has to be proper calibration of the robot to correct the joint variables and robot parameter, calibration of the camera to determine its accurate pose relative to world measurements, and then calibration of the hand-eye system to obtain the transformation of the camera relative to the robots hand.

REFERENCES

[1] G. Anzolin, A. Andreoni, and A. Zanfei, "Robot adoption and FDI driven transformation in the automotive industry," *Int. J. Automot. Technol. Manage.*, vol. 20, no. 2, pp. 215–237, 2020.

[2] E. Misimi, E. R. Øye, A. Eilertsen, J. R. Mathiassen, O. B. Åsebø, T. Gjerstad, J. Buljo, and Ø. Skotheim, "GRIBBOT—Robotic 3D vision-guided harvesting of chicken fillets," *Comput. Electron. Agricult.*, vol. 121, pp. 84–100, Feb. 2016.

[3] E. Bjorlykhaug and O. Egeland, "Vision system for quality assessment of robotic cleaning of fish processing plants using CNN," *IEEE Access*, vol. 7, pp. 71675–71685, 2019.

[4] M. Faisal, M. Alsulaiman, M. Arafah, and M. A. Mekhtiche, "IHDS: Intelligent harvesting decision system for date fruit based on maturity stage using deep learning and computer vision," *IEEE Access*, vol. 8, pp. 167985–167997, 2020.

[5] J. Wan, S. Tang, D. Li, M. Imran, C. Zhang, C. Liu, and Z. Pang, "Reconfigurable smart factory for drug packing in healthcare industry 4.0," *IEEE Trans. Ind. Informat.*, vol. 15, no. 1, pp. 507–516, Jan. 2019.

[6] M. Park, T.-A. Le, A. Eizad, and J. Yoon, "A novel shared guidance scheme for intelligent haptic interaction based swarm control of magnetic nanoparticles in blood vessels," *IEEE Access*, vol. 8, pp. 106714–106725, 2020.

[7] N. Jayaweera, P. Webb, and C. Johnson, "Measurement assisted robotic assembly of fabricated aero-engine components," *Assem. Autom.*, vol. 30, no. 1, pp. 56–65, Feb. 2010.

[8] Z. Li, S. Suntharasantic, S. Bai, and P. Chirarattananon, "Aeromechanic models for flapping-wing robots with passive hinges in the presence of frontal winds," *IEEE Access*, vol. 6, pp. 53890–53906, 2018.

[9] L. Wu, H. Li, Y. Li, and C. Li, "Position tracking control of tailsitter VTOL UAV with bounded thrust-vectoring propulsion system," *IEEE Access*, vol. 7, pp. 137054–137064, 2019.

[10] *International Federation of Robotics, IFR*. Accessed: Jan. 16, 2021. [Online]. Available: https://ifr.org/news/ifr-forecast-1.7-million-new-robots-to-transform-the-worlds-factories-by-20/

[11] M. Saedan and M. H. Ang, Jr., "3D vision-based control on an industrial robot," in *Proc. IASTED Int. Conf. Robot. Appl.*, 2001, pp. 152–157.

[12] N. Herakovic, *Robot Vision in Industrial Assembly and Quality Control Processes*. Rijeka, Croatia: InTech, 2010.

[13] E. N. Malamas, E. G. M. Petrakis, M. Zervakis, L. Petit, and J.-D. Legat, "A survey on industrial vision systems, applications and tools," *Image Vis. Comput.*, vol. 21, no. 2, pp. 171–188, Feb. 2003.

[14] S. Gobee, V. Durairajah, K. Xin, and L. L. Jie, "Robotic vision based PCB inspection with IOT interface," in *Proc. 3rd Int. Conf. Control, Robot. Cybern. (CRC)*, Sep. 2018, pp. 27–31.

[15] X. Fan, X. Wang, and Y. Xiao, "A combined 2D-3D vision system for automatic robot picking," in *Proc. Int. Conf. Adv. Mech. Syst.*, Aug. 2014, pp. 513–516.

[16] Y. Zhou, Q. Fang, K. Zhao, D. Tang, H. Zhou, G. Li, X. Xiang, and T. Hu, "Robust task-oriented markerless extrinsic calibration for robotic pick-and-place scenarios," *IEEE Access*, vol. 7, pp. 127932–127942, 2019.

[17] A. O. Fernandes, L. F. E. Moreira, and J. M. Mata, "Machine vision applications and development aspects," in *Proc. 9th IEEE Int. Conf. Control Autom. (ICCA)*, Dec. 2011, pp. 1274–1278.

[18] E. Martinez-Martin and A. P. D. Pobil, "Robot vision for manipulation: A trip to real-world applications," *IEEE Access*, vol. 9, pp. 3471–3481, 2021.

[19] Keyence. *2D Vision-Guided Robotics*. Accessed: Jan. 16, 2021. [Online]. Available: https://www.keyence.com/products/vision/vision-sys/2d-vgr/

[20] R. Kumar, S. Lal, S. Kumar, and P. Chand, "Object detection and recognition for a pick and place robot," in *Proc. Asia–Pacific World Congr. Comput. Sci. Eng.*, Nov. 2014, pp. 1–7.

[21] L. Mu, P. Yao, Y. Zheng, K. Chen, F. Wang, and N. Qi, "Research on SLAM algorithm of mobile robot based on the fusion of 2D LiDAR and depth camera," *IEEE Access*, vol. 8, pp. 157628–157642, 2020.

[22] B. Guo, H. Dai, Z. Li, and W. Huang, "Efficient planar surface-based 3D mapping method for mobile robots using stereo vision," *IEEE Access*, vol. 7, pp. 73593–73601, 2019.

[23] A. Shahzad, X. Gao, A. Yasin, K. Javed, and S. M. Anwar, "A vision-based path planning and object tracking framework for 6-DOF robotic manipulator," *IEEE Access*, vol. 8, pp. 203158–203167, 2020.

[24] S. Yang, S. Yang, and X. Yi, "An efficient spatial representation for path planning of ground robots in 3D environments," *IEEE Access*, vol. 6, pp. 41539–41550, 2018.

[25] M. Pena and R. Osorio, "Robot vision methodology for assembly manufacturing tasks," in *Proc. Electron., Robot. Automot. Mech. Conf. (CERMA)*, Sep. 2007, pp. 289–294.

[26] A. Pochyly, T. Kubela, V. Singule, and P. Cihak, "3D vision systems for industrial bin-picking applications," *Proc. Int. Conf. MECHATRONIKA*, 2012, pp. 1–6.

[27] J.-D. Lee, Y.-H. Wu, Y.-J. Jhao, L.-Y. Chen, and H.-I. Chen, "Development of mobile robot with vision inspection system and three-axis robot," in *Proc. 3rd Int. Conf. Control Robot. Eng. (ICCRE)*, Apr. 2018, pp. 6–10.

[28] Y. Kim, M.-H. Jeong, and D. J. Kang, "Mobile robot calibration," in *Proc. 13th Int. Conf. Control, Autom. Syst. (ICCAS)*, Oct. 2013, pp. 5504–5506.

[29] J.-W. Lee, G.-T. Park, J.-S. Shin, and J.-W. Woo, "Industrial robot calibration method using Denavit—Hatenberg parameters," in *Proc. 17th Int. Conf. Control, Autom. Syst. (ICCAS)*, Oct. 2017, pp. 1834–1837.

[30] J.-C. Hsiao, K. Shivam, I.-F. Lu, and T.-Y. Kam, "Positioning accuracy improvement of industrial robots considering configuration and payload effects via a hybrid calibration approach," *IEEE Access*, vol. 8, pp. 228992–229005, 2020.

[31] W. Qi, F. Li, and L. Zhenzhong, "Review on camera calibration," in *Proc. Chin. Control Decis. Conf.*, May 2010, pp. 3354–3358.

[32] X. Gong, Y. Lv, X. Xu, Z. Jiang, and Z. Sun, "High-precision calibration of omnidirectional camera using an iterative method," *IEEE Access*, vol. 7, pp. 152179–152186, 2019.

[33] Y. C. Shiu and S. Ahmad, "Calibration of wrist-mounted robotic sensors by solving homogeneous transform equations of the form AX=XB," *IEEE Trans. Robot. Autom.*, vol. 5, no. 1, pp. 16–29, Feb. 1989.

[34] N. Tan, X. Gu, and H. Ren, "Simultaneous robot-world, sensor-tip, and kinematics calibration of an underactuated robotic hand with soft fingers," *IEEE Access*, vol. 6, pp. 22705–22715, 2018.

[35] O. Faugeras, *Three-Dimensional Computer Vision: A Geometric Viewpoint*. Cambridge, MA, USA: MIT Press, 1993.

[36] A. Izaguirre, P. Pu, and J. Summers, "A new development in camera calibration: Calibrating a pair of mobile cameras," *Int. J. Robot. Res.*, vol. 6, no. 3, pp. 104–116, 1987.

[37] R. Y. Tsai and R. K. Lenz, "A new technique for fully autonomous and efficient 3D robotics hand/eye calibration," *IEEE Trans. Robot. Autom.*, vol. 5, no. 3, pp. 345–358, Jun. 1989.

[38] F. C. Park and B. J. Martin, "Robot sensor calibration: Solving AX=XB on the Euclidean group," *IEEE Trans. Robot. Autom.*, vol. 10, no. 5, pp. 717–721, Oct. 1994.

[39] J. C. K. Chou and M. Kamel, "Finding the position and orientation of a sensor on a robot manipulator using quaternions," *Int. J. Robot. Res.*, vol. 10, no. 3, pp. 240–254, Jun. 1991.

[40] G. Wei and K. G. A. Hirzinger, "Active self-calibration of robotic eyes and hand-eye relationships with model identification," *IEEE Trans. Robot. Autom.*, vol. 14, no. 1, pp. 158–166, Feb. 1998.

[41] S. Lee and S. Ro, "A self-calibration model for hand-eye systems with motion estimation," *Math. Comput. Model.*, vol. 24, nos. 5–6, pp. 49–77, Sep. 1996.

[42] J. Heller, M. Havlena, A. Sugimoto, and T. Pajdla, "Structure-from-motion based hand-eye calibration using L∞ minimization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2011, pp. 3497–3503.

[43] M. Ackerman and G. Chirikjian, "A probabilistic solution to the AX=XB problem: Sensor calibration without correspondence," in *Proc. Int. Conf. Geometric Sci. Inf.*, 2013, pp. 693–701.

[44] Q. Ma, H. Li, and G. S. Chirikjian, "New probabilistic approaches to the AX=XB hand-eye calibration without correspondence," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2016, pp. 4365–4371.

[45] J. Brookshire and S. Teller, "Extrinsic calibration from per-sensor egomotion," in *Robotics: Science and Systems VIII*. Cambridge, MA, USA: MIT Press, 2013, pp. 504–512.

[46] J. Brookshire and S. Teller, "Articulated pose estimation using tangent space approximations," *Int. J. Robot. Res.*, vol. 35, nos. 1–3, pp. 5–29, Jan. 2016.

[47] J. Brookshire and S. Teller, "Automatic calibration of multiple coplanar sensors," in *Robotics: Science and Systems VII*, vol. 33. Cambridge, MA, USA: MIT Press, 2012.

[48] N. Andreff, R. Horaud, and B. Espiau, "Robot hand-eye calibration using structure-from-motion," *Int. J. Robot. Res.*, vol. 20, no. 3, pp. 228–248, Mar. 2001.

[49] K. Pachtrachai, M. Allan, V. Pawar, S. Hailes, and D. Stoyanov, "Hand-eye calibration for robotic assisted minimally invasive surgery without a calibration object," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2016, pp. 2485–2491.

[50] X. X. Lu, "A review of solutions for perspective-n-point problem in camera pose estimation," *J. Phys., Conf. Ser.*, vol. 1087, Sep. 2018, Art. no. 052009.

[51] B. Zhou, Z. Chen, and Q. Liu, "An efficient solution to the perspective-n-point problem for camera with unknown focal length," *IEEE Access*, vol. 8, pp. 162838–162846, 2020.

[52] F. R. Gantmacher, *Matrix Theory*, vol. 21. New York, NY, USA: Chelsea, 1959, p. 6.

[53] R.-H. Liang and J.-F. Mao, "Hand-eye calibration with a new linear decomposition algorithm," *J. Zhejiang Univ.-Sci. A*, vol. 9, no. 10, pp. 1363–1368, Oct. 2008.

[54] H. Nguyen and Q. Pham, "On the covariance of X in $AX = XB$," *IEEE Trans. Robot.*, vol. 34, no. 6, pp. 1651–1658, Aug. 2018.

[55] Y.-C. Lu and J. C. K. Chou, "Eight-space quaternion approach for robotic hand-eye calibration," in *Proc. IEEE Int. Conf. Syst., Man Cybern. Intell. Syst. 21st Century*, Oct. 1995, pp. 3316–3321.

[56] H. H. Chen, "A screw motion approach to uniqueness analysis of head-eye geometry," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jan. 1991, pp. 145–146.

[57] K. Daniilidis and E. Bayro-Corrochano, "The dual quaternion approach to hand-eye calibration," in *Proc. 13th Int. Conf. Pattern Recognit.*, 1996, pp. 318–322.

[58] N. Andreff, R. Horaud, and B. Espiau, "On-line hand-eye calibration," in *Proc. 2nd Int. Conf. 3-D Digit. Imag. Modeling*, 1999, pp. 430–436.

[59] D. Condurache and A. Burlacu, "Orthogonal dual tensor method for solving the AX=XB sensor calibration problem," *Mechanism Mach. Theory*, vol. 104, pp. 382–404, Oct. 2016.

[60] D. Condurache and I.-A. Ciureanu, "A novel solution for AX=YB sensor calibration problem using dual lie algebra," in *Proc. Int. Conf. Control, Decis. Inf. Technol.*, 2019, pp. 302–307.

[61] S. Gwak, J. Kim, and F. C. Park, "Numerical optimization on the Euclidean group with applications to camera calibration," *IEEE Trans. Robot. Autom.*, vol. 19, no. 1, pp. 65–74, Feb. 2003.

[62] J. Heller, D. Henrion, and T. Pajdla, "Hand-eye and robot-world calibration by global polynomial optimization," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2014, pp. 3157–3164.

[63] Z. Zhao, "Simultaneous robot-world and hand-eye calibration by the alternative linear programming," *Pattern Recognit. Lett.*, vol. 127, pp. 174–180, Nov. 2019.

[64] Z. Zhang, L. Zhang, and G.-Z. Yang, "A computationally efficient method for hand–eye calibration," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 12, no. 10, pp. 1775–1787, 2017.

[65] J. Wu, Y. Sun, M. Wang, and M. Liu, "Hand-eye calibration: 4-D procrustes analysis approach," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 6, pp. 2966–2981, Jun. 2020.

[66] K. Strobl and G. Hirzinger, "Optimal hand-eye calibration," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2006, pp. 4647–4653.

[67] Z. Roth, B. Mooring, and B. Ravani, "An overview of robot calibration," *IEEE J. Robot. Autom.*, vol. JRA-3, no. 5, pp. 377–385, Oct. 1987.

[68] X. Shi and Y. Liu, "Reprojection error based annealed particle filter for human upper body pose reconstruction," in *Proc. IEEE Int. Conf. Service Oper. Logistics, Informat.*, Oct. 2014, pp. 277–281.

[69] I. Ali, O. J. Suominen, E. R. Morales, and A. Gotchev, "Multi-view camera pose estimation for robotic arm manipulation," *IEEE Access*, vol. 8, pp. 174305–174316, 2020.

[70] W. Yin, S. Pathak, A. Moro, A. Yamashita, and H. Asama, "Accurate all-round 3D measurement using trinocular spherical stereo via weighted reprojection error minimization," in *Proc. IEEE Int. Symp. Multimedia (ISM)*, Dec. 2019, pp. 86–867.

[71] R. Zhu, H. K. Galoogahi, C. Wang, and S. Lucey, "Rethinking reprojection: Closing the loop for pose-aware shape reconstruction from a single image," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 57–65.

[72] J. A. Hesch and S. I. Roumeliotis, "A direct least-squares (DLS) method for PnP," in *Proc. Int. Conf. Comput. Vis.*, Nov. 2011, pp. 383–390.

[73] W. Wang, G.-J. Zhang, and Z.-Z. Wei, "A framework for 2D model-based pose estimation," in *Proc. Int. Conf. Comput. Sci. Appl. (CSA)*, Nov. 2015, pp. 48–51.

[74] S. Zhang, H. Yu, J. Dong, T. Wang, L. Qi, and H. Liu, "Combining Kinect and PnP for camera pose estimation," in *Proc. 8th Int. Conf. Hum. Syst. Interact. (HSI)*, Jun. 2015, pp. 357–361.

[75] *Thermobot: Autonomous Robotic System for Thermo-Graphic Detection of Cracks*. Accessed: Apr. 2, 2021. [Online]. Available: https://thermobot.eu

[76] K. Koide and E. Menegatti, "General hand–eye calibration based on reprojection error minimization," *IEEE Robot. Autom. Lett.*, vol. 4, no. 2, pp. 1021–1028, Apr. 2019.

[77] V. M. Eskov, V. F. Pyatin, V. V. Eskov, and L. K. Ilyashenko, "The heuristic work of the brain and artificial neural networks," *Biophysics*, vol. 64, no. 2, pp. 293–299, Mar. 2019.

[78] O. I. Abiodun, M. U. Kiru, A. Jantan, A. E. Omolara, K. V. Dada, A. M. Umar, O. U. Linus, H. Arshad, A. A. Kazaure, and U. Gana, "Comprehensive review of artificial neural network applications to pattern recognition," *IEEE Access*, vol. 7, pp. 158820–158846, 2019.

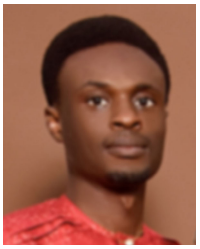
[79] M. Mainampati and B. Chandrasekaran, "Evolution of machine learning algorithms on autonomous robots," in *Proc. 10th Annu. Comput. Commun. Workshop Conf. (CCWC)*, Jan. 2020, pp. 737–741.

[80] Y. Guo, X. Ke, J. Ma, and J. Zhang, "A pipeline neural network for low-light image enhancement," *IEEE Access*, vol. 7, pp. 13737–13744, 2019.

[81] A. Avelar, I. Salgado, H. Ahmed, M. Mera, and I. Chairez, "Differential neural networks observer for second order systems with sampled and quantized output," *IFAC-PapersOnLine*, vol. 51, no. 13, pp. 490–495, 2018.

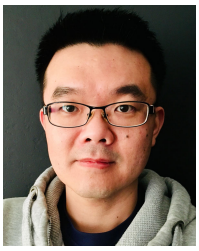
[82] I. Salgado, H. Ahmed, O. Camacho, and I. Chairez, "Adaptive sliding-mode observer for second order discrete-time MIMO nonlinear systems based on recurrent neural-networks," *Int. J. Mach. Learn. Cybern.*, vol. 10, no. 10, pp. 2851–2866, Oct. 2019.

[83] A. Guarneros, I. Salgado, M. Mera, and H. Ahmed, "Differential neural network identifier with composite learning laws for uncertain nonlinear systems," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 7897–7902, 2020.
[84] J. Hua and L. Zeng, "Hand–eye calibration algorithm based on an optimized neural network," *Actuators*, vol. 10, no. 4, p. 85, Apr. 2021.
[85] M. Mishra and M. Srivastava, "A view of artificial neural network," in *Proc. Int. Conf. Adv. Eng. Technol. Res. (ICAETR)*, Aug. 2014, pp. 1–3.
[86] C. Zhang and W. Xu, "Neural networks: Efficient implementations and applications," in *Proc. IEEE 12th Int. Conf. ASIC (ASICON)*, Oct. 2017, pp. 1029–1032.
[87] D. K. Chaturvedi, *Soft Computing: Studies in Computational Intelligence*. Berlin, Germany: Springer, 2008.
[88] I. Bilbao and J. Bilbao, "Overfitting problem and the over-training in the era of data: Particularly for artificial neural networks," in *Proc. 8th Int. Conf. Intell. Comput. Inf. Syst. (ICICIS)*, Dec. 2017, pp. 173–177.
[89] A. P. Piotrowski and J. J. Napiorkowski, "A comparison of methods to avoid overfitting in neural networks training in the case of catchment runoff modelling," *J. Hydrol.*, vol. 476, pp. 97–111, Jan. 2013.
[90] J. Salvi, X. Armangué, and J. Batlle, "A comparative review of camera calibrating methods with accuracy evaluation," *Pattern Recognit.*, vol. 35, no. 7, pp. 1617–1635, Jul. 2002.
[91] J. Mallon and P. F. Whelan, "Which pattern? Biasing aspects of planar calibration patterns and detection methods," *Pattern Recognit. Lett.*, vol. 28, no. 8, pp. 921–930, Jun. 2007.
[92] NI Vision. *Perspective and Nonlinear Distortion Calibration*. Accessed: May 22, 2021. [Online]. Available: https://documentation.help/NI-Vision-LabView-Basics/Perspective_and_Nonlinear_Distortion_Calibration.html
[93] Z. Zhang, "Flexible camera calibration by viewing a plane from unknown orientations," in *Proc. 7th IEEE Int. Conf. Comput. Vis.*, Sep. 1999, pp. 666–673.
[94] J. Heikkila and O. Silvén, "A four-step camera calibration procedure with implicit image correction," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 1997, pp. 1106–1112.
[95] P. F. Sturm and S. J. Maybank, "On plane-based camera calibration: A general algorithm, singularities, applications," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 1999, pp. 432–437.
[96] L. Gao, Y. Fang, Z. Lin, and K. Chen, "A versatile camera calibration technique for high accuracy 3D machine vision metrology," *Mech. Sci. Technol.*, vol. 5, pp. 323–344, Aug. 1987.
[97] Z. Wang, Z. Wang, and Y. Wu, "Recognition of corners of planar checkboard calibration pattern image," in *Proc. Chin. Control Decis. Conf.*, May 2010, pp. 3224–3228.
[98] Longer Vision Technology. *Camera Calibration Using a Circle Grid*. Accessed: May 22, 2021. [Online]. Available: https://longervision.github.io/2017/03/18/ComputerVision/OpenCV/opencv-internal-calibration-circle-grid/
[99] R. Jain, R. Kasturi, and B. G. Schunck, *Machine Vision*, vol. 5. New York, NY, USA: McGraw-Hill, 1995.
[100] D. Hu, D. DeTone, and T. Malisiewicz, "Deep ChArUco: Dark ChArUco marker pose estimation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 8436–8444.
[101] T. Clarke and X. Wang, "The control of a robot end-effector using photogrammetry," *Int. Arch. Photogramm. Remote Sens.*, vol. 33, no. B5/1, pp. 137–142, 2000.
[102] I. Rabinowitz, J. Broomberg, M. Goitein, K. McCarthy, and J. Leong, "Accuracy of radiation field alignment in clinical practice," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 11, no. 10, pp. 1857–1867, 1985.
[103] J. Wu, M. Liu, C. Zhang, and Z. Zhou, "Correspondence matching and time delay estimation for hand-eye calibration," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 10, pp. 8304–8313, Apr. 2020.
[104] F. Zhong, Z. Wang, W. Chen, K. He, Y. Wang, and Y.-H. Liu, "Hand-eye calibration of surgical instrument for robotic surgery using interactive manipulation," *IEEE Robot. Autom. Lett.*, vol. 5, no. 2, pp. 1540–1547, Apr. 2020.
[105] M. Zhou, M. Hamad, J. Weiss, A. Eslami, K. Huang, M. Maier, C. P. Lohmann, N. Navab, A. Knoll, and M. A. Nasseri, "Towards robotic eye surgery: Marker-free, online hand-eye calibration using optical coherence tomography images," *IEEE Robot. Autom. Lett.*, vol. 3, no. 4, pp. 3944–3951, Oct. 2018.
[106] R. Bruder, F. Griese, F. Ernst, and A. Schweikard, "High-accuracy ultrasound target localization for hand-eye calibration between optical tracking systems and three-dimensional ultrasound," *Bildverarbeitung für die Medizin*. Berlin, Germany: Springer, 2011, pp. 179–183.
[107] H. Li, Q. Ma, T. Wang, and G. S. Chirikjian, "Simultaneous hand-eye and robot-world calibration by solving the $AX = YB$ problem without correspondence," *IEEE Robot. Autom. Lett.*, vol. 1, no. 1, pp. 145–152, Dec. 2015.
[108] K. Pachtrachai, F. Vasconcelos, G. Dwyer, V. Pawar, S. Hailes, and D. Stoyanov, "CHESS—Calibrating the hand-eye matrix with screw constraints and synchronization," *IEEE Robot. Autom. Lett.*, vol. 3, no. 3, pp. 2000–2007, Jul. 2018.
[109] G. Yang, W. Haixia, X. Peng, and C. Maoyong, "Development of high-precise hand-eye calibration software system," in *Proc. 27th Chin. Control Conf.*, Jul. 2008, pp. 426–429.
[110] C.-H. Wu, "The kinematic error model for the design of robot manipulator," in *Proc. Amer. Control Conf.*, Jun. 1983, pp. 497–502.
[111] P.-N. Le and H.-J. Kang, "Robot manipulator calibration using a model based identification technique and a neural network with the teaching learning-based optimization," *IEEE Access*, vol. 8, pp. 105447–105454, 2020.
[112] N. Aghakhani, M. Geravand, N. Shahriari, M. Vendittelli, and G. Oriolo, "Task control with remote center of motion constraint for minimally invasive robotic surgery," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2013, pp. 5807–5812.
[113] Y. K. Gu, W. F. Li, J. Zhang, and G. Q. Qiu, "Effects of wear, backlash, and bearing clearance on dynamic characteristics of a spur gear system," *IEEE Access*, vol. 7, pp. 117639–117651, 2019.
[114] H. Zhang, W. Qin, Y. Gao, Q. Li, Z. Chen, and J. Zhao, "Disturbance elimination for the modular joint torque sensor of a collaborative robot," *Math. Problems Eng.*, vol. 2020, pp. 1–14, Oct. 2020.
[115] A. Olabi, M. Damak, R. Bearee, O. Gibaru, and S. Leleu, "Improving the accuracy of industrial robots by offline compensation of joints errors," in *Proc. IEEE Int. Conf. Ind. Technol.*, Mar. 2012, pp. 492–497.
[116] S. J. D. Prince, *Computer Vision: Models Learning and Inference*. Cambridge, U.K.: Cambridge Univ. Press, 2012.
[117] G. Dwyer, F. Chadebecq, M. T. Amo, C. Bergeles, E. Maneas, V. Pawar, E. V. Poorten, J. Deprest, S. Ourselin, P. De Coppi, T. Vercauteren, and D. Stoyanov, "A continuum robot and control interface for surgical assist in fetoscopic interventions," *IEEE Robot. Autom. Lett.*, vol. 2, no. 3, pp. 1656–1663, Jul. 2017.
[118] X. Li, C. Sun, W. Cheng, X. Jiang, and Y.-H. Liu, "Adaptive vision-based control for robotic tiling with uncalibrated cameras and limited FOV," in *Proc. IEEE 15th Int. Conf. Control Autom. (ICCA)*, Jul. 2019, pp. 168–173.
[119] V. Villani, F. Pini, F. Leali, and C. Secchi, "Survey on human–robot collaboration in industrial settings: Safety, intuitive interfaces and applications," *Mechatronics*, vol. 55, pp. 248–266, Nov. 2018.
[120] S. Robla-Gómez, V. M. Becerra, J. R. Llata, E. González-Sarabia, C. Torre-Ferrero, and J. Pérez-Oria, "Working together: A review on safe human-robot collaboration in industrial environments," *IEEE Access*, vol. 5, pp. 26754–26773, 2017.
[121] F. Schramm, F. Geffard, G. Morel, and A. Micaelli, "Calibration free image point path planning simultaneously ensuring visibility and controlling camera path," in *Proc. IEEE Int. Conf. Robot. Autom.*, Apr. 2007, pp. 2074–2079.
[122] R. Fareh, M. Baziyad, T. Rabie, and M. Bettayeb, "Enhancing path quality of real-time path planning algorithms for mobile robots: A sequential linear paths approach," *IEEE Access*, vol. 8, pp. 167090–167104, 2020.



IKENNA ENEBUSE received the B.Eng. degree in electrical and electronics engineering from the University of Benin, Nigeria, in 2012, and the M.Sc. degree in systems and control engineering from Coventry University, U.K., in 2014, where he is currently pursuing the Ph.D. degree under the joint collaboration between Coventry University and Oxford Vision and Sensor Technology (OVST).

His research interests include machine vision, robotics, real-time control systems, and sensor fusion.



MATHIAS FOO (Member, IEEE) received the B.Eng. (Hons.) and M.Eng.Sc. degrees in electronic engineering from Multimedia University, Malaysia, in 2002 and 2004, respectively, and the Ph.D. degree from the University of Melbourne, Melbourne, VIC, Australia, in 2012. He is currently a Senior Teaching Fellow with the University of Warwick, Coventry, U.K. Prior to this appointment, he was a Postdoctoral Research Fellow with the Asia Pacific Center for Theoretical Physics (APCTP), Pohang, South Korea, from 2012 to 2015, a Research Fellow with the Warwick Integrative Synthetic Biology Centre, School of Engineering, University of Warwick, from 2015 to 2018, and an Assistant Professor with the School of Mechanical, Aerospace and Automotive Engineering, Coventry University, U.K., from 2018 to 2021. His research interests include dynamical system modeling, and application of control system to industrial, medical, and biological problems.

BABUL SALAM KSM KADER IBRAHIM (Member, IEEE) received the Ph.D. degree from The University of Sheffield, U.K., in 2011. He is currently an Assistant Professor of automation and robotic with the School of Mechanical, Aerospace and Automotive, Coventry University, U.K. Prior to this appointment, he was an Associate Professor with the Department of Mechatronic and Robotic Engineering, University Tun Hussein Onn Malaysia, Johor, Malaysia. He is a highly motivated academic with significant expertise in the field mechatronics with special adherence to robotics and control systems with 12 years teaching experience combined with course work and research background. He has published more than 100 technical articles in journals and conference proceedings in these fields. His research interests include electric vehicle, mechatronics, nonlinear control and modeling, rehabilitation robotics, and control in biomedical application. He is a Chartered Engineer of the Institution of Engineering and Technology (IET), U.K.

HAFIZ AHMED (Senior Member, IEEE) received the Ph.D. degree in automatic control from the University of Lille 1, Villeneuve-d'Ascq, France, in 2016.

From 2016 to 2021, he was with Clemson University, Clemson, SC, USA, Asia Pacific University, Dhaka, Bangladesh, North South University, Dhaka, Coventry University, Coventry, U.K., and Birmingham City University, Birmingham, U.K. Since April 2021, he has been with the Nuclear Futures Institute, Bangor University, U.K., where he is currently the Sêr Cymru Senior Lecturer (Associate Professor) in nuclear instrumentation and control. His research interest includes applied control engineering with special focus in energy and environment. He was a recipient of the European Embedded Control Institute (EECI) Ph.D. Award, in 2017, and the Best Ph.D. Theses Award from the Research Cluster on Modelling, Analysis, and Management of Dynamic Systems (GDR-MACS) of the National Council of Scientific Research (CNRS) in France, in 2017. He is an Associate Editor of the *International Journal of Electrical Engineering and Education*. He is also actively involved in organizing special issues and sessions at various international journals and conferences.

FHON SUPMAK received the B.Sc. degree in control engineering and the M.Sc. degree in electronic and computer based system design.

She is currently the Chief Executive Officer with the Oxford Vision and Sensor Technology (OVST), she is responsible for guiding the commercial and technical strategy of the company. She has over 15 years' experience in 2D and 3D machine vision system development. She has delivered several successful 3D vision system projects internationally for companies, including Jaguar Land Rover and Aston Martin, Ford, Caterpillar, and Siemens. She also works her role as an Industrial Advisory Board (IAB) with Coventry University, U.K., and King Mongkut's Institute of Technology Ladkrabang (KMITL), Thailand. OVST specializes in the design and manufacture of 2D and 3D machine vision systems. OVST develops 3D sensors that incorporate intellectual property developed at Oxford University. It also received the British Government's 3D Vision Technology Development Award and the Henry Ford Technical Achievement Award.

ODONGO STEVEN EYOBU received the B.Sc. degree in computer science from Islamic University, Uganda, in 2004, the M.Sc. degree in data communication and software engineering from Makerere University, Uganda, in 2007, and the Ph.D. degree in electronics engineering from Kyungpook National University, South Korea, in 2018. He is currently a Lecturer with the School of Computing and Informatics Technology, Makerere University. His research interests include deep learning systems, indoor localization, vehicular communications, intelligent transportation systems, and wireless sensors.