



A survey on vision guided robotic systems with intelligent control strategies for autonomous tasks

Abhilasha Singh, V. Kalaichelvi & R. Karthikeyan |

To cite this article: Abhilasha Singh, V. Kalaichelvi & R. Karthikeyan | (2022) A survey on vision guided robotic systems with intelligent control strategies for autonomous tasks, Cogent Engineering, 9:1, 2050020, DOI: [10.1080/23311916.2022.2050020](https://doi.org/10.1080/23311916.2022.2050020)

To link to this article: <https://doi.org/10.1080/23311916.2022.2050020>



© 2022 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.



Published online: 03 Apr 2022.



Submit your article to this journal 



Article views: 7452



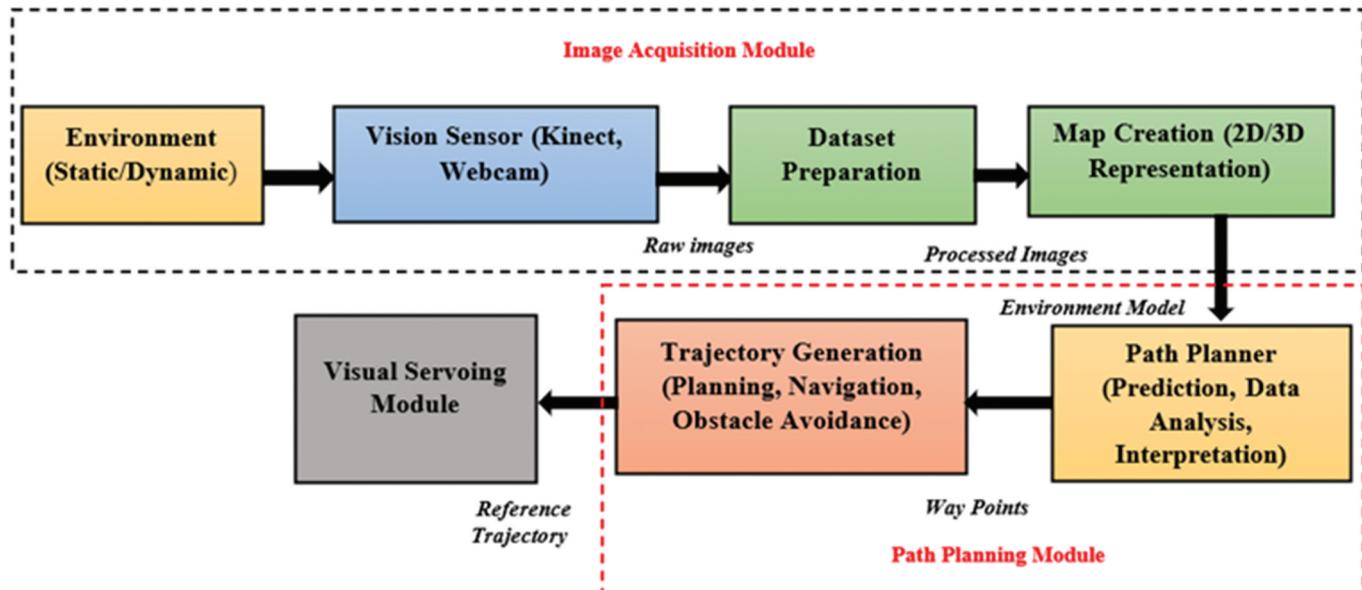
View related articles 



View Crossmark data 



Citing articles: 13 View citing articles 



A survey on vision guided robotic systems with intelligent control strategies for autonomous tasks

Abhilasha Singh, V. Kalaichelvi and R. Karthikeyan

Cogent Engineering (2022), 9: 2050020



Received: 04 August 2021

Accepted: 01 March 2022

Corresponding author: Abhilasha Singh, Electrical and Electronics Engineering, Birla Institute of Technology and Science Pilani - Dubai Campus, United Arab Emirates
E-mail: p20180906@dubai.bits-pilani.ac.in

Reviewing editor:
Ahmed Haj Darwish, Faculty of Informatics Engineering, University of Aleppo, Aleppo SYRIAN ARABREPUBLIC

Additional information is available at the end of the article

SYSTEMS & CONTROL | REVIEW ARTICLE

A survey on vision guided robotic systems with intelligent control strategies for autonomous tasks

Abhilasha Singh¹, V. Kalaichelvi² and R. Karthikeyan²

Abstract: The Vision Guided Robotic systems (VGR) is an essential aspect of modern intelligent robotics. The VGR is rapidly transforming manufacturing processes by enabling robots to be highly adaptable and intelligent reducing the cost and complexity. For any sensor-based intelligent robots, vision-based planning is considered as one of the most prominent steps followed by controlled actions using visual feedback. To develop robust vision-based autonomous systems in robotic applications, path-planning and localization can be implemented along with Visual Servoing (VS) for robust feedback control. In the available literature, most of the reviews are focused on a particular module of autonomous systems like path planning, motion planning strategies, or Visual Servoing techniques. In this paper overall review of different modules in vision-guided robotic systems is presented. So, this review provides researchers with broader in-depth knowledge about different modules that exist in the vision-guided autonomous system. The review also includes different vision sensors that are commonly used in industries covering their characteristics and applications. In this work, overall, 227 research papers in path planning and vision-based control algorithms are reviewed with recent intelligent

ABOUT THE AUTHOR



Abhilasha Singh

Abhilasha Singh is currently pursuing full time Ph.D. degree in Electrical and Electronics engineering at BITS Pilani, Dubai Campus since 2018. Her current research area focuses on Robotics, Computer Vision, and Intelligent Control Techniques.

PUBLIC INTEREST STATEMENT

Nowadays, robots are considered as an important element in society with the fact that they can replace humans in basic and dangerous activities. Among them, Vision-Guided Robots (VGR) are an essential aspect of modern intelligent robotics that can easily navigate in any environment using visual feedback thereby increasing productivity and efficiency. The authors have highlighted the important layers that exist in any vision-guided robot and focus on state-of-art techniques and developments that took place in the last 5 years. The goal of this review paper was to make researchers understand the basic concepts of vision-guided robots and what are the current gaps existing in this field. The evolution of research in vision-guided robots have been illustrated through graphical analysis. Finally, the future scope of this study has also been highlighted in the paper. The VGR can help in transforming manufacturing processes highly adaptable reducing the cost and complexity. The research activities specified are

techniques based on optimization and learning-based approaches. The graphical analysis illustrating the advancements of research in the field of vision-based robotics using Artificial Intelligence (AI) is also discussed. Lastly, this paper concludes by discussing some of the research gaps, challenges, and future directions existing in vision-based planning and control.

Subjects: Artificial Intelligence; Robotics; Systems & Control Engineering

Keywords: path planning; visual servoing; vision guided robotic systems; vision sensors; control techniques; artificial intelligence

1. Introduction

Nowadays, robots are considered an important element in society with the fact that they can replace humans in basic and dangerous activities. Among them, vision-guided robots are commonly used in industries that can easily navigate in any environment using the feedback obtained from the vision sensor. These robots increase productivity making the robots highly adaptable and robust. There are different modules to build a vision-guided system namely perception, localization, path planning, and control. As far as robot planning is concerned it is difficult for a robot to react to sudden environmental changes or avoid any kind of obstacles whereas humans can easily achieve these tasks (Pandey et al., 2017). In robot planning, a sequence of actions is planned from start to goal point by using planning algorithms simultaneously avoiding obstacles (Garrido et al., 2011; Hardlick & Shapiro, 1993). However, designing an efficient navigation strategy is the most important issue in creating intelligent robots. Therefore, the robot planning problem using vision sensors is one of the most interesting and wide research areas where the ultimate goal is to achieve a safe and optimal route for robot navigation. In general, vision-based robotic systems can be applied to several industrial applications like spray painting (Ferreira et al., 2014), pick and place, assembly task in the optical industry, automotive, robotic welding like pipe and spot welding (Rout et al., 2019), payload identification and much more where efficient planning and control are of prime importance. The modules involved in vision-based autonomous robotic systems are illustrated in Figure 1.

1.1. Perception

In robotics, the perception stage involves visualization and reaction to the changes in the surrounding environment based on the features obtained from the vision sensor which is used for making decisions, planning, and operation in real-world environments. The key modules of a robot perception system are essentially sensor data acquisition and data pre-processing using image processing so that an environment model can be generated. Some of the applications of robotic perception are obstacle detection (Rout et al., 2019), object recognition (Bore et al., 2018), semantic segmentation (Sunderhauf et al., 2016), 3D environment representation (Saarinen et al., 2012), gesture and voice recognition

Figure 1. Process flow diagram of vision-based autonomous robots.

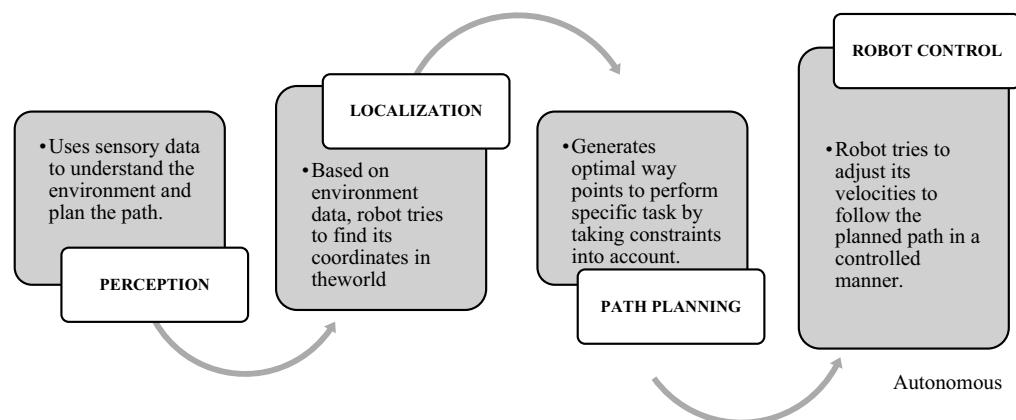
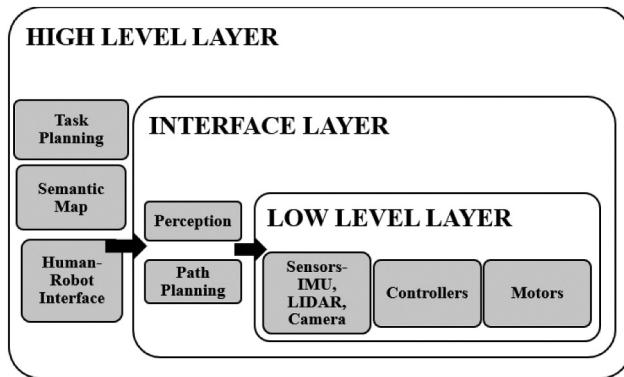


Figure 2. Layers of vision-based autonomous robots.



(Fong et al., 2003), activity classification (Faria et al., 2015), terrain classification (Manduchi et al., 2005), road detection (Fernandes et al., 2014), vehicle detection (Asvadi et al., 2017).

1.2. Localization

Localization refers to finding the position and orientation of the robot in space relative to the camera (Paya et al., 2017). In other words, the robot maps out the world while keeping track of where it is in the world using bounding boxes or markers around the region of interest in an image or video sequence. The applications of localization are robot vehicle tracking, pedestrian detection (Premebida & Nunes, 2013), human detection (Sarabu & Santra, 2021), and environment change detection (Guo et al., 2020).

1.3. Path planning

The main idea behind vision-based robot planning (Mezouar & Chaumette, 2002) is to generate optimal waypoints along with visual constraints. Once the waypoints are obtained, trajectories with respect to time can be generated so that the robot can be made to move along the planned trajectories making the error between the reference and actual trajectories becomes zero. Path planning is commonly used in autonomous vehicles, mobile robot navigation, tracing weld lines in the welding industry, video games, and industrial automation in logistics (Dolgov et al., 2008). The path planning algorithms can be designed for both static and dynamic environments where the optimal path is obtained with the knowledge of any object positions in that environment map because most of the obstacles in real-life situations are dynamic. The path planning is categorized based on robot workspace, obstacle, and environment which is discussed in detail in section 2.3.

1.4. Robot control

Once the path is planned, the joints of the robot are controlled to achieve the desired goals which are also termed trajectory tracking. For the design of control techniques, the mathematical model of the robot is obtained using kinematics, inverse kinematic, or robot dynamics. There are many control algorithms to track the desired trajectory namely conventional controllers like Proportional Integral Derivative (PID) controller, Adaptive Control schemes (Dan Z et al., 2017), Optimal Control and Robust Control (Mittal et al., 2003; Nagrath et al., 2003) followed by intelligent controllers like Neural Network, Fuzzy Logic, Genetic Algorithm (Arhin et al., 2020). This paper provides a detailed survey of state-of-art optimization and learning-based approaches for designing the controllers which is one of the contributions of this paper. Figure 2 illustrates the different layers of any vision-based autonomous robotic system. There are three layers—a high level, an interface layer, and a low-level layer which constitutes autonomous robotic systems. This paper mainly concentrates on vision sensor-based robotic systems.

The communication of any robotic system is from a high-level to low-level layer in which task planned or any human-machine commands are sent to the interface layer which acts as

middleware between robot hardware and environment. Based on the action received from the environment, the interface layer decides the path and obstacles. This planned path is sent to sensors installed on the robot which in turn controls the wheel velocities and move accordingly. Though the robots have become intelligent with the advancement of technologies, some flaws still exist in vision-based systems particularly if the camera movement is dynamic and in highly illumination environment. These factors hinder the planning stage thereby misleading the robots. In recent times, many reviews and research have been reported in vision based robotic systems. Pandey et al. (2017) discussed in depth survey on intelligent techniques for navigation of mobile robots. The author focussed on evolutionary algorithms, deterministic and non-deterministic algorithms whereas Rout et al. (2019) explained different sensors and techniques to track the weld seam in robotic welding applications. The author also explained different techniques for determining weld seam position and geometry feature extraction along with advantages and disadvantages. Mouli et al. (2013) presented review on embedded systems for object sorting using vision-guided robotic arm. The author addressed past and current design methodologies, image processing techniques, modeling of arm and finally wireless control of robotic arms for sorting applications. Halme et al. (2018) discussed recent method in vision-based technologies for human-robot interaction. The focus in this paper was vision-based safety and alarming systems based on flexibility, system speed and collaboration level. The author also addressed challenges existing in human-robot collaboration along with future direction. Fauadi et al. (2018) discussed recent intelligent techniques for mobile robot navigation and the author focussed three vision systems namely monocular, stereo and trinocular system. Chopde et al. (2017) discussed recent developments in vision-based system for agricultural inspection along with image processing techniques. Zafar and Mohanta (2018) presented commonly used techniques focusing on conventional and heuristic approaches for path planning of mobile robots whereas Campbell et al. (2020) discussed various global and local path planning algorithms for both static and dynamic environments. Gul et al. (2021) present consolidated review on path planning and optimization techniques for ground, aerial and under water vehicles along with communication strategies. Wang, Fang, et al. (2020) focussed on different Visual Servoing strategies along with trajectory planning methods. The author discussed trajectory planning methods based on potential field and optimization algorithms. Dewi et al. (2018) presented a review on Visual Servoing design and control strategies for agricultural robots whereas Huang et al. (2019) discussed Visual Servoing techniques for non-holonomic robots for two different robot-camera configurations namely independent and coupling types. The author proposed active disturbance rejection control (ADRC) technique for implementing Visual Servoing with the uncalibrated camera.

From the above discussion, it can be noticed that most of the present literature is related to path planning, motion planning, calibration strategies, and Visual Servoing (VS) techniques. There is a limited consolidated review on vision-based intelligent robots with current state-of-art techniques. Hence, in this review, an attempt has been made to present a detailed survey on different modules focussing on different layers like perception, localization, path planning, and control strategies in vision-based robotic systems. This paper also elaborates on recent techniques discussed in various literatures for vision-based robotic systems. This paper also provides an in-depth review of sensors commonly used for real-time applications, different calibration procedures, image processing techniques for feature extraction, and Visual Servoing algorithms for robot control. Apart from this, high-level controller design based on recent developments has been addressed along with future direction and challenges. The objectives of the current review are to focus on the following:

- To discuss the different modules of the perception stage that covers major areas like different sensors used in real-time applications, different camera calibration techniques, image processing techniques and addressing their advantages and disadvantages.
- To elaborate on different localization methods and its advancements made in recent years.
- To present some of the state-of-art intelligent techniques for vision-based path planning which can be used to produce waypoints to plan its task.

- To discuss different Visual Servoing strategies along with limitations of each approach.
- To discuss different high-level optimization and learning-based controllers for effective visual feedback control.

The paper is divided into six sections. The description of different modules of vision-based robot planning and types of Visual Servoing is discussed in Section 2. It is followed by an explanation of a high-level controller design for Visual Servoing in section 3. The graphical analysis of the literature review for this paperwork is discussed in section 4. In Section 5, the research gap, challenges, and future scope of vision-based planning are discussed followed by a conclusion in Section 6.

2. Modules of vision-based autonomous robot

This paper concentrates on path planning and control using vision sensor since it is widely used in industrial applications. The vision-based robot planning is an amalgamation of pre-processed images in the perception module which is further needed for planning the robot path to perform sequential actions in a controlled manner as illustrated in Figure 3. The detailed explanation for each module of vision-based robot planning and control will be explained in further sections.

2.1. Perception module for vision-based robot planning

In this section, different elements in the perception stage are discussed which covers vision sensors for acquiring data, camera calibration methods, and finally image processing for feature extraction. The main purpose of this module is to pre-process the visual information so that the robot can understand the environment. Based on the information, an environmental representation or map is created which helps the robot to decide on its motion. The vision sensors used for robotic perception system comprises lighting source, lenses, image sensors, controllers, vision tools, and communication protocols (D-M. Lee et al., 2004). The internal hardware structure of the vision sensor is shown in Figure 4.

Figure 3. Modules of vision-based autonomous robot.



Figure 4. Internal hardware structure of vision sensors used for robot perception system.

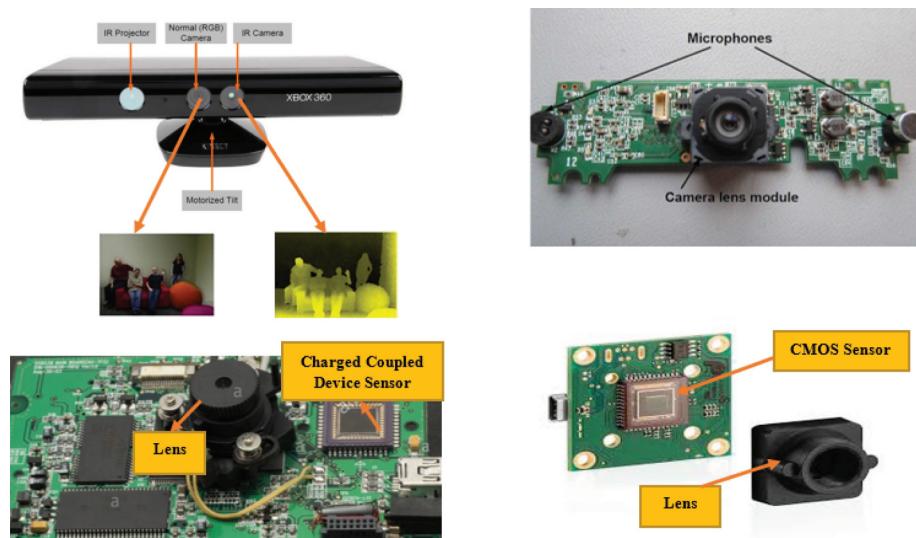
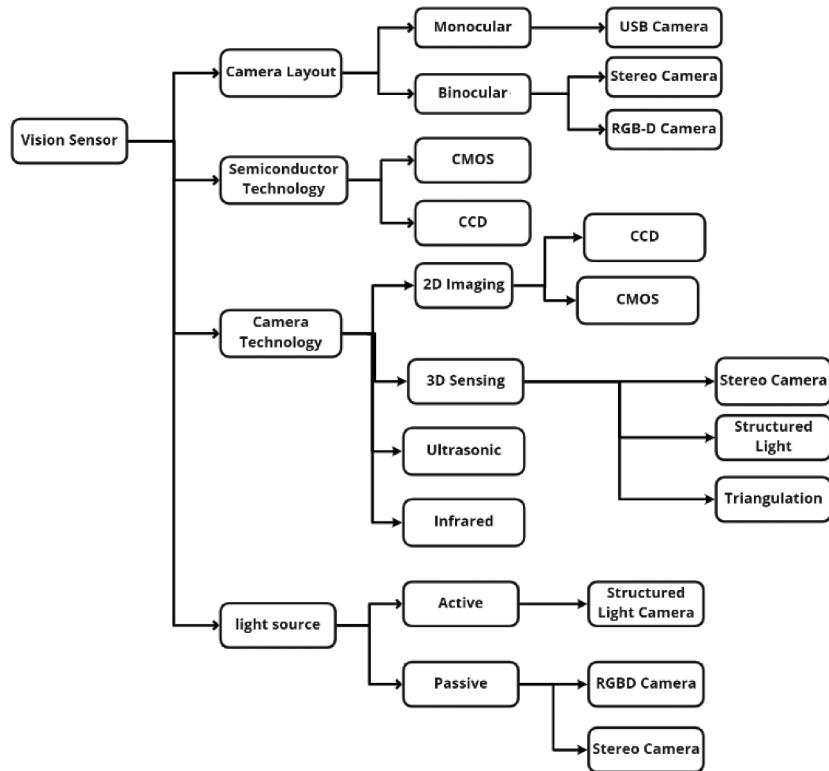


Figure 5. Classification of different vision sensors used for robot perception system.



The basic working of the vision sensor is that initially, the lens performs object inspection by looking for specific features to determine position, orientation, and accuracy. Once the field of view (FOV) is set, image sensors convert light into a digital image which is sent to the Analog to Digital (ADC) converter for processing and optimizing images. In general, sensors are classified as exteroceptive and proprioceptive where exteroceptive sensors measure external parameters like distance from the robot and proprioceptive measure internal parameters of the robot—speed and wheel rotations. The robots are always installed with exteroceptive sensors to correct the errors. Since the camera belongs to the exteroceptive sensor, its detailed classification based on the camera layout, semiconductor technology, camera technology, and based on the light source used are illustrated in Figure 5. Environment, lighting, and modularity are some of the most important factors to be considered while selecting a vision sensor (Malamas et al., 2003) for specific applications. Apart from this other factors like application, sensor type, and communication interfaces should be considered for choosing the right camera. The most common communication interfaces used are USB 2.0, USB 3.0, GigE, Camera Link, PoE, and VGA. Vision sensors have enormous applications like food inspection, welding, object tracking, auto sorting the waste, assistive robots, assembly applications, and much more (Nandini et al., 2016). In recent years, there are enormous companies that manufacture smart vision sensors for various industrial applications with advanced technologies. They have built-in high-end processors which help in faster processing of frames per second. Table 1 explains the main categories of vision sensors with detailed specifications like manufacturer, Frames Per Second(FPS). It also gives readers a better idea of choosing a particular sensor based on applications.

2.1.1. Camera calibration methodology

Any sensors which are installed on robots need to be calibrated to avoid measurement errors. The commonly used sensor in robots is the camera module which has a lens in form of a pinhole or fisheye. The raw image obtained from the camera is not the real view hence some transformation has to be applied so that it looks the same way as the ideal camera captures. This transformation

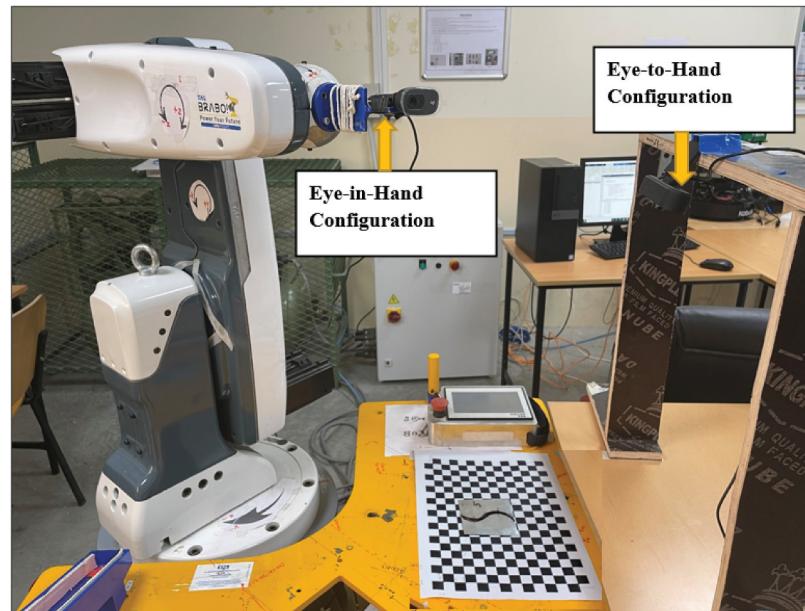
Table 1. List of vision sensors used for robot planning

S. No	Sensor Type	Manufacturer	Resolution	Frame Rate(fps)	Advantages	Disadvantages	Applications
1	2D Sensing	Logitech, Robotiq wrist camera	1920x1080	30	Less expensive, easy to integrate and highly reliable.	Cannot operate fully autonomously and poor working under low light conditions	Self-Localization, Spy Robots, object detection
2	RGBD Camera	Microsoft Kinect, Asus Xtion sensor, Orbec Astra	RGB = 640x480 Depth = 1280x1024	15-30	Adaptable to complex varying work conditions, highly flexible, increases productivity.	More expensive, need proper system like GPUs for point cloud processing	3D Sensing of Environment, Human Recognition, Gesture Recognition
3	Stereo	Zed 2 Ensenso, Tara, Intel Realsense D435	3840x1080, 2560x720, 4416x1242, 1280 x 720.	15-100	Simulate human binocular vision, works better for long distance and moving objects, works better in outdoor applications.	Needs powerful processor, CPU intensive due to two cameras.	Mobile Robot Navigation, Human Tracking, Gesture Recognition, 3D Surface Visualization
4	CCD	Fisnar F4000-AKIT-VISION, ikon-XL Large CCD	795x596, 4096x4112	20	Good Pixel to pixel reproducibility, Has high quality ADC	Expensive, Slow to readout, Needs more circuitry.	Robot Mapping, Machine Vision Applications
5	CMOS	Samsung, Omniprecision, Canon, Panasonic.	2592x1944	14-53	Less expensive, Faster read speed.	12-bit ADC reduces the image quality, Linearity and sensitivity variations are high.	Virtual Reality, Machine Vision, SLAM Applications
6	Ultrasonic Camera	Fluke Acoustic camera i900	1280x800	12.5	Longer and wider range, not affected by light conditions dust, colors, transparency, or the reflective properties or surface texture of the object.	Poor in defining edges, affected by temperature, humidity, and pressure.	good for measuring liquid levels and distances, detecting objects, and avoiding collisions on land or under water.
7	Infrared Camera	Samsara, FLIR 380x, FLIR MR265	160x120, 1080p/720p	9	Good in capturing edges, Cost effective	Sensitive to IR light and sunlight	Identify humans, objects in low light conditions

is obtained by computing the intrinsic parameters and the distortion parameters (D' & Gasbarro, 2017). The first and foremost step is to calibrate the camera to remove lens distortion and calculate camera internal and external parameters. The camera calibration parameters include intrinsic parameters namely focal length, lens distortion, and principal point whereas extrinsic parameters namely rotation and translation of the camera. Based on the lens model, calibration is performed for two main models namely the pinhole model (Rigoberto et al., 2020; Wang, Zhang et al., 2019) and the fish-eye model (Lee, 2020). Since general explanations have been already covered in various literature, this section explains the comparison of calibration methods for different lens types. The camera for calibration is commonly placed in the eye-in-hand or eye-to-hand configuration as shown in Figure 6. If the camera is placed in an eye-to-hand setup, a checkboard can be placed either in the robot workspace or it can be attached to the robot end-effector. Furthermore, if the camera is placed in an eye-in-hand setup, a calibration board can be placed in the robot workspace or placed somewhere outside the workspace which can be visualized by the camera. Based on the different camera placement, transformation calculations vary.

The general classification of camera calibration methods is conventional visual calibration, camera self-calibration, and active vision-based calibration (Long & Dongri, 2019). In traditional visual calibration, intrinsic parameters and extrinsic parameters are computed. The intrinsic parameters like focal length, skew, the principal point of the lens is constant whereas external parameters like translational and rotation components are dynamic. Traditional calibration methods include the DLT algorithm, Tsai calibration method (Tsai, 1987), linear transformation method and RAC-based calibration Method, Zhang Zhengyou calibration method based on the chessboard (Zhang, 2000). The most commonly and widely used technique for calibration is Zhang's Method. In this technique, checkboard pattern images are taken at different angles using a pinhole camera. Next, the 4×3 camera matrix is obtained which is the combination of the intrinsic and extrinsic matrices. Using these matrices, a camera to world transformation can be found by mapping the checkboard corner points and world corner points. The world $[XYZ]$ to camera coordinates $[X_cY_cZ_c]$ are transformed using extrinsic matrix whereas camera $[X_cY_cZ_c]$ to pixel coordinates $[uv]$ are obtained using an intrinsic matrix. The intrinsic matrix I and extrinsic matrix E of the pinhole model is given by

Figure 6. Eye-in-hand and eye-to-hand configuration for performing camera calibration.



$$\mathbf{I} = \begin{bmatrix} f_x & 0 & 0 \\ S & f_y & 0 \\ c_x & c_y & 1 \end{bmatrix}; \mathbf{E} = \begin{bmatrix} R_{3 \times 3} \\ t_{1 \times 3} \end{bmatrix} \quad (1)$$

Where (f_x, f_y) = Focal length in pixels

(c_x, c_y) = principal point of lens in pixels

S = Skew coefficient

The camera matrix is given by,

$$\mathbf{C}_{4 \times 3} = \mathbf{E}_{4 \times 3} * \mathbf{I}_{3 \times 3} \quad (2)$$

Finally using the camera matrix image points (u, v) are converted to corresponding world coordinates (X, Y, Z) which is given by,

$$w[u \ v \ 1]_{1 \times 3} = [X \ Y \ Z \ 1]_{1 \times 4} \mathbf{C}_{4 \times 3} \quad (3)$$

Where w is the scaling factor and the above matrices do not account for distortions so to represent the real camera radial and tangential distortion has to be accounted for. The detailed explanations and derivations can be referred to in (Zhang, 2000). Furthermore, camera self-calibration involves finding the unknown camera intrinsic parameters from a sequence of images and it does not require any special calibration objects. The commonly used self-calibration techniques are Absolute Conic (Richard Hartley & Chang, 1992) and Absolute Quadric (Trigg, 1997). Next in active vision-based calibration, image data is obtained from a controlled camera, and camera motion is known during the calibration process. Active vision calibration methods are further classified as Makonde's three orthogonal translational motion methods (Trigg, 1997) and Hartley camera pure rotation correction method (Hartley, 1997). The summary of related research works of different methodologies and recent developments in calibration technologies is discussed in detail in Table 2.

The advantages and disadvantages of different camera calibration methods are listed in Table 3.

2.1.2. Dataset preparation for vision-based robot planning

In a robotic vision-based task, the next important step after calibration is the preparation of the dataset captured from vision sensors for processing. Once the images are obtained from sensors, the raw image has to be processed using some image processing algorithms to extract useful information. In vision-based path planning, the dataset is nothing but the collection of 2D or 3D images from which the robot has to understand the data and locomote in a particular fashion. In general, 2D image can be defined by two-dimensional function $f(x, y)$ where x and y are height and width which range from 0 to 255 (Bovik, 2009d) and f is the intensity of the image at that point whereas 3D image has an additional depth which contributes to finding the object distance easily. Hence many upcoming research works are focussing on 3D based Vision tasks which creates realistic effects easily.

The common classification of images storage formats is given in Figure 7.

In the case of binary images, it has two-pixel elements 0 and 1 where 0 refers to black and 1 refers to white. In 8-bit format, images are also called the grayscale image has 256 colors where 0 stands for Black, 255 stands for white, and 127 stands for gray. Finally, the 16-bit color format also called RGB format has 65,536 different colors in it which are divided into three channels namely Red, Green, and Blue (Jackson, 2016). Based on these image formats, there are different image processing techniques to extract the useful features and prepare the datasets suitable to the algorithm so that robots are not distorted. The common process flow of data processing to extract

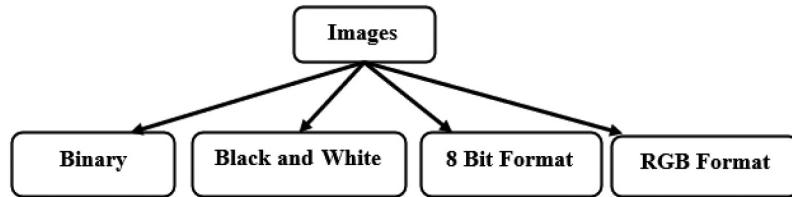
Table 2. Summary of recent works collected in camera calibration methodology

S. No	Reference	Calibration Method	Camera	Marker used	Reprojection Error (Pixels)	Remarks
1	Gunen et al. (2020)	Zhang and Bouguet with Evolutionary algorithms	Sony Alpha A6000 camera	Checkerboard	-	Zhang method performed better than Bouguet with mean value of 0.0041.
2	Chen et.al (2020)	Zhang with Digital Image Correlation algorithm(DIC)	GRAS-50SSM-C, Point Grey Research, Inc.; Schneider—Kreuznach Xenoplan 50 mm lens	Synthetic random speckle pattern, circular and checkerboard pattern	Noiseless image =0.004 Real image =0.038	higher control point match accuracy, and precise estimation of camera parameters
3	Huang et al. (2020)	Camera self-calibration	Skybotix stereo VI-sensor	EUROC dataset	-	0.142° and 0.032 m errors in orientation and position for camera-IMU calibration
4	Pan et al. (2021)	Camera self-calibration with quantum genetic algorithm and feature matching	Multi color camera UI-2240RE-C of IDS company	No checkboard was used	0.15 mm	Calibration can be performed for any real time applications without special boards.
5	Pirahansiah et al. (2015)	Zhang method with image quality assessment	-	checkboard	1.76-0.98 (pixels)	Mean projection errors were calculated for different number of images

Table 3. Comparative analysis of different camera calibration methods

S. No	Calibration Method	PROS	Cons	Calibration Toolbox	Complexity	Quality Metrics	Checkboard
1	Zhang method	1.NEEDS THREE DIFFERENT VIEWS OF PLANAR PATTERN 2. HIGH ACCURACY 3. OFFLINE PROCESS	1. Needs high quality images for accurate calibration 2. Needs knowledge of calibration process	Zhang	low	high	yes
2	Tsai Method	1. PLANAR AND NON-PLANAR PATTERNS CAN BE USED. 2.TWO-STEP CALIBRATION PROCESS	Skew and distortion is not into account	Heikkila	low	high	Yes
3	Batista Method	1.NEEDS ONLY ONE IMAGE FOR CALIBRATION 2. ITERATIVE METHOD WITH LEAST SQUARE TECHNIQUE	Distortion is modelled using one parameter	Computer vision, Bouquet	Low	high	yes
4	Active Vision Based	REQUIRES IMAGE IN CONTROL ENVIRONMENTS	Medium accuracy is only possible	Camera calibration (MATLAB)	Medium	low	No (only frames)
5	Self-calibration	PRE PROCESSING OF IMAGES NOT REQUIRED	Low accuracy with high computational time	Camera calibration (MATLAB)	High	low	No (Many Images)

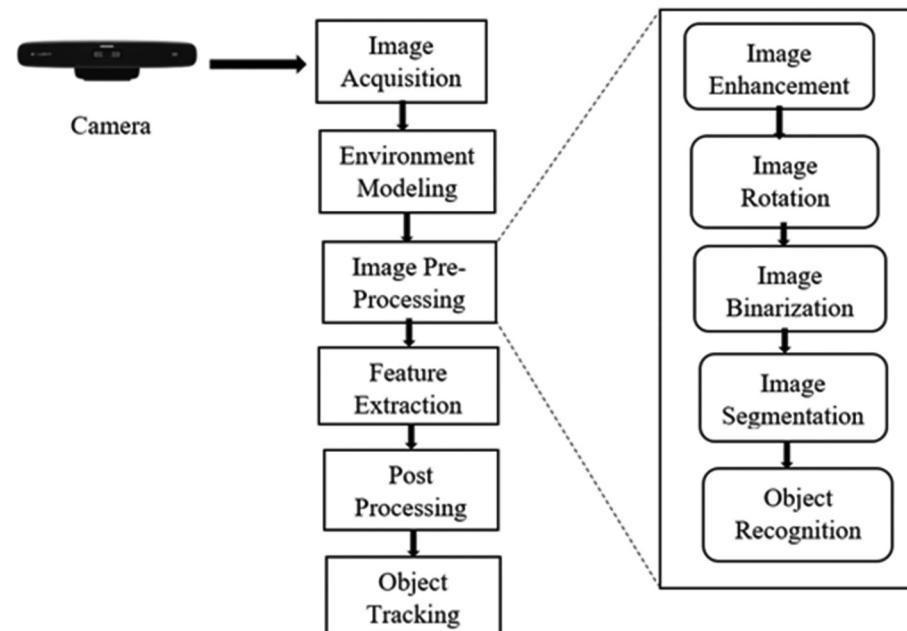
Figure 7. Classification of different types image storage formats.



the features from the image acquired from the vision sensor is shown in Figure 8. The process involves image acquisition from a vision sensor followed by an image preprocessing step where the images are scaled, rotated, binarized, and segmented to extract the useful features (Patel & Bhavsar, 2021; Premebida et al., 2018). Once the preprocessing is done, the datasets can be used for any kind of applications like object tracking, motion estimation, object recognition and so on (Garibotto et al., 2013; Song et al., 2018). Though the process sounds simple there are a lot of factors that affect the processing stage. The most common factors are high illumination and noise in images. Also with high-resolution cameras, large datasets need a lot of memory and powerful processing devices. Working with large videos poses serious problems where better encryption and compression techniques are needed. To address these issues, still many research works are being conducted.

In recent times, there are many improvements made over conventional image processing techniques with the advancements of artificial intelligence algorithms. Byambasuren et al. (2020) applied image processing techniques for conducting quality assurance of products using ABB IRB 120 industrial arm. The author captured 3D data of normal and abnormal objects and used HSV transformation with histogram normalization technique to distinguish the objects. The main gap in that paper is that it was not tested in real-time considering illumination effects and other noises. Cruz et al (2020) used the structured light camera for welding inspection of LPG vessels. The weld lines were extracted using image processing techniques like grayscale conversion and binarization. The author was able to achieve a quality index of 95%-99% in practical validation. S Xu et al. (2020) presented a

Figure 8. Steps involved in preparation of dataset from vision sensor for robot path planning.



review on computer vision and image processing techniques for construction sites since monitoring the process is difficult whereas Vijay Kakani et al. (2020) reviewed different computer vision methods and artificial techniques for the food industry. The author also discusses the possibility and future directions of industry 4.0 for sustainable food production. Rodriguez Martinez et al. (2020) presented a vision-based navigation system using a photometric A star algorithm that compares previous states (images) with the current one to create an efficient visual path.

In recent years, there has seen a massive increase in the adoption of computer vision tools and software in various applications like camera calibration (Hua & Zeng, 2021), feature detection/tracking (Wang Zhang et al., 2020), image positioning for sorting (Man Li et al., 2020), speech classification (Mustafa et al., 2020), segmentation (L. C. Santos et al., 2020), thresholding (Angani et al., 2020), quality assurance/monitoring (Byambasuren et al., 2020) and 3D image processing (Cui et al., 2020). Numerous toolkits, frameworks, and software libraries have been developed which make the system intelligent and smart. Some of the best open-source computer vision tools which can be used for creating an effective computer vision application are mentioned in Table 4 with specific languages supported (Bebis et al., 2003; Parikh, 2020). This helps the readers to pick the right tools based on software compatibility and specific applications.

Finally, once the datasets are prepared and features are extracted, the sensor measurements are used to create a map to represent the environment. The map data is then sent to the localization module to find the robot's position in the real world. From the above discussions, it can be inferred that the perception module is the key element in vision-based robotic systems because without this module robot cannot understand its surroundings. In-depth discussions of different elements of perception modules like sensors, calibration techniques, and image processing techniques are covered with state-of-art developments made by the researchers. The future of perception is that it can be used in Industry 4.0 which makes the industrial process more fluid and flexible. Since industries are looking for a reduction in energy consumption and recycling capabilities, in such cases artificial intelligent techniques and image processing helps in improving the quality of production and enhancing the automation control.

2.2. Localization module for vision-based robot planning

Localization is a process of estimating the position and orientation of the robot using sensor information. There are two main types of localization namely indoor localization and outdoor localization. The indoor localization is used for position estimation when there is no GPS information available whereas the outdoor localization is based on a satellite navigation system. Based on position information, it can be further classified as position tracking, global localization, and kidnapped problem (Aguiar et al., 2020). In position tracking, the robot's initial position is known, and the objective is to track the robot at each instance of time using the previous position obtained from odometry, IMU sensor, and vision sensor. In case of large position uncertainty, position tracking does not perform well. Hence global localization is chosen where the robot does not have any knowledge about its initial position. Furthermore, in some cases, the robot is moved to an unidentified place which is called the kidnapped problem. Hence an autonomous robot should be able to monitor its location and relocation simultaneously by applying the relocation method. To address this issue, there are many localization algorithms- probabilistic localization strategies, automatic map-based localization, RFID-based approaches, and evolutionary techniques (Liu et al., 2007). This section explains the recent works based on indoor localization methods using vision sensor in which images of different visual markers like QR codes, AprilTags, ARTags are used as the input (Morar et al., 2020). The summary of works implemented in localization modules is discussed in Table 5.

The above table describes various works that can be performed in the field of robot localization in both indoor and outdoor environments. Many sensors were reported in the literature works and most commonly IMU sensors, LiDAR, and cameras are used, and they can also be fused using different filtering techniques. The common filters that are used for sensor fusion are Kalman filter, 1D complementary filter, and EKF filter. Other hybrid approaches using LSTM, machine learning

Table 4. Common computer vision tools for robotic perception system

S.No	Tools	Purpose	Software Interface	URL	Pros	Cons
1	OpenCV	Computer vision applications and accelerate machine perception	C, P, M, J	opencv.org	1. Open source 2. Users can modify the algorithms.	Not easy to use like MATLAB®
2	BoofCV	Real-time robotics and Computer vision applications	J	boofcv.org	1. User-friendly interface 2. Provides multiple language support	Slower performance for low-level operations like motion analysis, optical flow computation, and image matching.
3	TensorFlow	Open-source library for machine learning and artificial intelligence	C, P, M, J	https://www.tensorflow.org/	1. Open-source 2. Compatible with multiple programming languages	Needs more resources and toolkits
4	PyTorch	Optimized tensor library primarily used for Deep Learning applications using GPUs and CPUs	C,P,J	https://pytorch.org/	1. Support for CPU and GPU 2.Powerful API with computational graph support	1. Does not support visualization tools like Tensorboard. 2. Developer community is small.
5	Computer vision Toolbox	Designing and testing computer vision, 3D vision, and video processing systems.	C, P, M	https://www.mathworks.com/products/image.html	1. Large images can be processed easily. 2. Many pre-trained models are available. 3. Can accelerate algorithms by GPUs.	1. It is not an open-source platform 2. Performance is comparatively slower for real-time video processing.
6	CUDA	Provides GPU-accelerated image, video, and signal processing functions	C, P, M	https://developer.nvidia.com/cuda-downloads	1. Multiple language support. 2. Fast and effective	High power consumption
7	Keras	Interface for machine learning platform TensorFlow	C, P, M, J	https://keras.io/	1. Easy to use, fast and user friendly 2.mutiple backend support	Debugging is difficult and some features needs improvement

(Continued)

Table 4. (Continued)

S.No	Tools	Purpose	Software Interface	URL	Pros	Cons
8	OpenVINO	Optimization of deep learning models and deployment of AI into hardware	C, P	https://www.intel.com/content/www/us/en/developer/tools/openvino-toolkit/overview.html	1.open-source toolkit 2. Multiple deep learning models are available	Limited support for python
9	ViSP	Used to compute homography, pose, Visual Servoing and camera calibration	C, P	visp.inria.fr	1. Simple and highly portable. 2. Has predefined 2D and 2D features	1. multiple platforms are not supported as of now.
10	Meshlab	Processing 3D data	J, C	www.meshlab.net	1.open-source platform 2. Supports different input/output file formats like .stl, .obj, .collada	Some tools takes longer time to process so not exactly user friendly.
11	VISO suite	Computer vision platform for business applications like industrial automation, remote inspection and visual monitoring	C, P, cloud-based SaaS solution	https://viso.ai/	1.Provides robust and safe AI deployment 2.Cost and time effective 3.Cross platform migration is possible	Not an open-source platform

C—C/C++; P—Python; M—MATLAB; J—Java

Table 5. Common computer vision tools for robotic perception system

S. No	Reference	Methods	Markers Used	Sensor	Environment used
1	Al-Forati and Roshid (2018)	Binary search algorithm and minimum bounded circle algorithm	LED arrays	Light Dependant Resistor (LDR)	Indoor
2	Lai and Su (2018)	Filter SLAM and Graph SLAM	RGBD map	Microsoft Kinect Camera	Indoor
3	Xu shmaliy et.al (2019)	Unbiased and Extended FIR filters; Kalman filters	NA	LiDAR with compass	Indoor
4	Xu shmaliy et.al (2019)	Faster R-CNN,LSTM and AMCL	Images	Monocular camera with Hokuyo LiDAR	Indoor
5	Almeida et al. (2018)	Bayes classifier, Multilayer Perceptron (MLP), k-Nearest Neighbors KNN, Support Vector Machines (SVM) and Optimum Path Forest (OPF)	NA	Ultrasonic Sensor array	Indoor
6	Neto et al. (2009)	Complementary Filter	NA	IMU, GPS, digital compass and odometry sensor	Outdoor
7	Qi Li et al. (2020)	(2D) Map matching localization approach based on Delaunay triangulation.	3D LiDAR map	LiDAR	Outdoor
8	Ali Durmush et al. (2020)	Visual odometry using ORB SLAM	FinnForest dataset	Four RGB cameras, an IMU, and a GNSS receiver.	Outdoor
9	Reis et al. (2019)	Lidar Odometry and Mapping in Real-time (LOAM)	Forest dataset	3D laser data, thermal camera, inertial units, GNSS, and RGB camera	Outdoor

Table 6. Comparison of common robot localization techniques

S. No	Localization Methods	Advantages	Disadvantages	Applications
1	Probabilistic Methods	1. Can be used to globally localize the robot. 2. Highly accurate	Difficult to localize with large maps	Indoor navigation
2	Automatic map based	1. Precise localization is possible using prior maps. 2. Highly robust and reliable as they send information to the robot rapidly	Memory requirement is high particularly for 3D maps.	Urban planning and outdoor navigation
3	RFID based method	1. Cost effective 2. Easy maintenance 3. Suitable for industrial environments	1. Exhaustive localization not possible 2. Range of detection is less than 1 m	Industrial Selective object identification and navigation
4	Evolutionary technique	1. Highly accurate than probabilistic methods. 2. Does not require noise distribution. 3. Efficient tracking in complex environments	1. Requires more sensors for localization 2. Errors due to slip and attitude is not accounted	Location routing problems in logistics

techniques can be used to find the efficient and accurate position estimates of the robot. The advantages and disadvantages of different localization methods along with applications are elaborated and compared in Table 6.

This section discussed in detail the recent techniques which are used for efficient localization strategy as it is the most important step in robot planning. Without proper localization, robot position cannot be estimated and might lead to a kidnapped problem. There are many works reported on how to recover the robot from being lost in an unknown environment (Hu et al., 2020; Le et al., 2021). In such scenarios reinforcement learning-based localization is being researched.

2.3. Vision-based path planning

In general, Robot planning is the problem of deciding what the robot should try to do under constraints and uncertainties caused due to mapping, localization estimates, and the outcome of actions. The planning is generally classified into task planning, path planning, and motion planning. Path planning is the extension of localization which is the problem of finding a safe and optimal route in cartesian space or joint space. Next figuring out how to move between planned paths by computing velocity and acceleration is termed motion planning (Patle et al., 2019; Tzafestas Spyros, 2018). In this paper, a detailed survey is done for path planning using vision sensors followed by motion planning in terms of Visual Servoing. The task planning is not the scope of this paper. The information obtained from the perception section is used by the robot to plan the predefined path for moving in its world coordinates. Path planning is generally classified into local and global based on obstacle type. Local path planning is performed by obtaining data from vision sensors simultaneously when the robot is moving changes in the environment. In global path planning, paths are generated only if the environment is known with the start and endpoints (Jixin et al., 2019). Path planning can be implemented in both mobile robots and manipulators to perform a specific task but in mobile robots, they have a higher degree of freedom because of the moving base, hence path planning of such

robots is more challenging than serial manipulators (Gasparetto et al., 2015). Once the path for a robot is described, the next step is to generate the trajectory which serves as the input to the motion control system ensuring that the robot follows the executed trajectory. Different techniques like Cubic polynomial, Quintic polynomial, and Beizer curves are used are commonly used for trajectory generation. Some of the factors to find the optimal route for path planning of robotic manipulators and mobile robot navigation includes optimization on acceleration, velocity, and jerk (Lin, 2014; Tang et al., 2012). Once the trajectory is generated, the robot must move in these trajectories efficiently by controlling its pose which is taken care by the Visual Servoing module. The overall process of vision-based planning is illustrated in Figure 9.

The current methods of robot planning specified in the literature are categorized into conventional methods that include Roadmap, Cell Decomposition, and Potential Fields followed by intelligent approaches which include genetic algorithm, Neural Network (NN), fuzzy logic, and artificial intelligence techniques (Chaumette et al., 2008). The detailed path planning algorithms, challenges, types can be found in a few works of literature (Campbell et al., 2020; Jiang & Ma, 2020; Nadhir Ab Wahab et al., 2020; LC Santos et al., 2020; H. Y. Zhang et al., 2018). The path planning methodologies varies based on the workspace of the robot involving 2D or 3D data. The research dated back involves planning mostly in a 2D environment since it was easy for the robot to understand and also it was easy to process the data. But in order to make the analysis more realistic to the real-time environment, 3D data can be used for path planning which is termed 3D path planning. With the advancement in vision and artificial intelligence, processing and handling of 3D data became much easier hence recent research concentrates more on 3D data analysis. Another aspect of classification is based on environment model details which can be known or unknown. In most cases, environment details are known, and path planning is performed but in the case of an unknown environment, a Reinforcement learning approach can be implemented. Another classification considered is based on obstacles. The detailed classification of different path planning algorithms used in the research is illustrated in Figure 10

In recent years, metaheuristic algorithms for path planning like a genetic algorithm (GA) is commonly used as it helps in the optimization of several factors like path length, path time, path smoothening (Goldberg, 1989). Another intelligent technique used for path planning is the neural network where paths generated are fed as inputs to the neurons. As it moves from input layers to output layers via hidden layers weights and biases are updated. To improve the training of the neural network, hidden layers can be increased, and also momentum coefficient can be specified for faster convergence. Artificial Intelligent (AI) algorithms can be implemented when humans are unaware of process knowledge, handle complex processes, and process huge datasets. Machine learning and deep learning are a subset of AI techniques which has widespread use in applications

Figure 9. Process flow diagram of vision-based path planning.

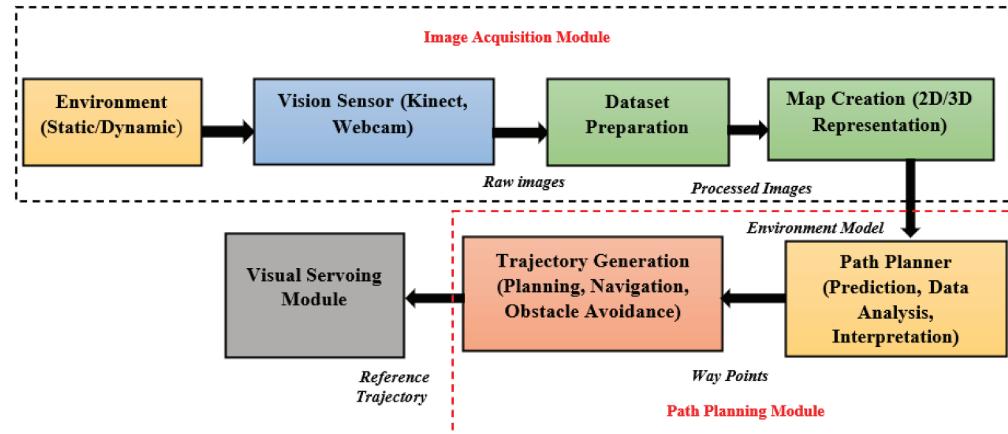
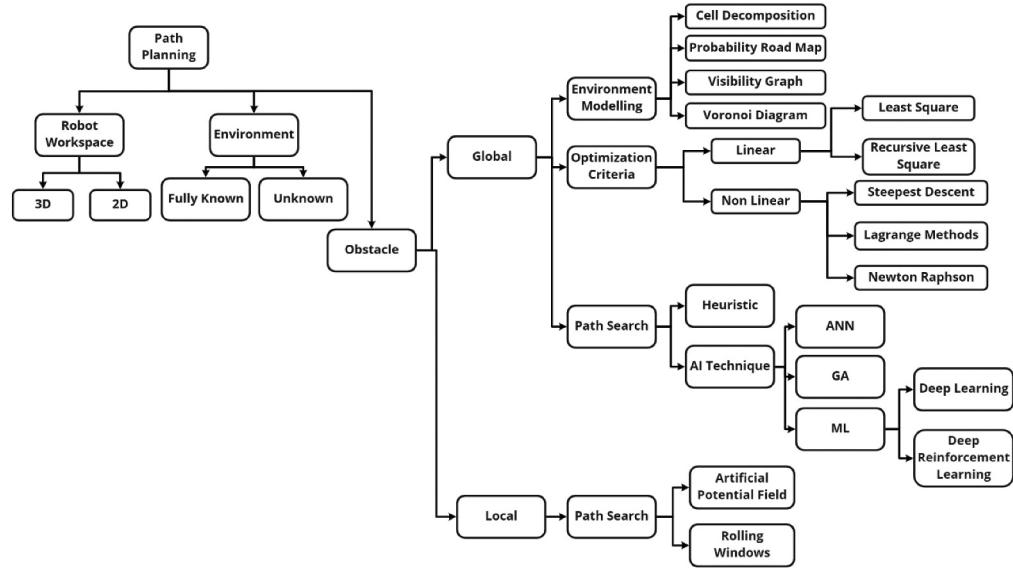


Figure 10. Different algorithms for vision-based path planning.



like image classification and object recognition, predicting grasp locations, bin-picking operations, trajectory planning, motion estimation, pose estimation, camera delocalization, remote sensing using Convolutional Neural Network (CNN) (Alom et al., 2019; Bhatt et al., 2021). Another new direction has evolved in path planning where the algorithm learns by itself. Based on the actions and reward obtained, it defines the path which is called Reinforcement Learning (RL) (Wang Liu et al., 2020). In RL, a path is generated based on learning from mistakes. If the training is repeatedly done, learning will be better. Also combining deep learning and RL is a popular area of research in recent times. The current research trends in vision-based path planning algorithms based on the papers surveyed are listed in Table 7.

The recent developments in the field of vision-based path planning were discussed in the table above. It can be seen that many works have been reported in deep learning and reinforcement learning. The main advantage of such methods is that they can be used if the robot does not have any information about the environment thereby there is no need for complex mathematical modeling. The conventional vision-based path planning needed additional feature extraction algorithms and they were susceptible to odometry noise. With deep learning networks, feature extraction is done automatically without the need for additional feature extraction algorithms. They can easily handle larger datasets without additional processing. Nowadays, 3D path planning is being implemented where the environment is uncertain like underwater, forest, and urban areas (Yang et al., 2016). Unlike 2D path planning complexity of 3D path planning is high when kinematic constraints increases. Hence several solutions are being investigated to address this issue.

2.4. Visual servoing techniques for robot navigation

As per the process flow explained in section 2.4, the waypoints and trajectories generated as fed as inputs to the Visual Servoing module where robots are made to move along the planned trajectories in a controlled manner. Visual Servoing in general controls the robot motion by using visual information as feedback (Ke et al., 2017). In general, using image processing (Wang Fang et al., 2020), computer vision (Kanellakis et al., 2017), robotics (Zake et al., 2020), and control theory (Sarapura et al., 2018) visual feedback control can be achieved. The path planner module can be combined with Visual Servoing for robust feedback control because the path planning algorithms take into account critical constraints and uncertainties in the system (Kazemi et al., 2010). The two main types of Visual Servoing are Position-Based Visual Servoing (PBVS) and Image-based Visual

Table 7. Literature review of vision-based planning tasks applied to manipulators and mobile robots

S. No	References	Applications	Algorithms	Sensing Device	Performance Metrics	Manipulator used
1	Ali Gong et al (2020); Dai et al. (2019); Su et al. (2021)	Path planning	Ant colony and Markov Decision Process (MDP); Ant colony with A* algorithm	No	Average path length and time; convergence speed; shortest path length and bending suppression effect	Simulated mobile robot
	Doulamis et al. (2018); Al-Jarrah et al. (2018)	Fuzzy based Edge detection with GA based A*	2D camera		Precision; Path length and computational time	4-wheel robot with manipulator; 4 wheel mobile robot
	Qingyang et al. (2012); L. Gong et al. (2021)	SVM	Lidar; velodyne Lidar with IMU		Path smoothness; curvature analysis; Time	4-wheel vehicle
	Wu et al. (2017); Chen Sandev et al (2017)	DCNN-based fractional order	Camera; Stereo camera		Validation accuracy; Average cycle time	Mobile robot, Pioneer 3AT, mobile base with 3DOF soft gripper
	Luo et al. (2015); Fereh et al. (2020)	Lazy PRM algorithm with RRT; PRM	Xtion PRO depth camera; RGBD camera		Path length, path efficiency, trackability and swiftness	KUKA IIWA Wheeled mobile robots; Pioneer P3-DX
	W. He et al. (2019); Trigatti et al. (2018)	5 th order Biezer Curve	Nil		Velocity, torque and acceleration; error	Robot GR 680; 3DOF industrial manipulator, Mobile Robot
2	Nuzzi et al. (2019)	Tip position control	CNN	CMOS camera	Translational precision; rotational error; average elapsed time	Universal robots UR10
	G Li et al. (2019)		SIFT, SURF-based feature extractor	Four Genius Widecam F100 cameras; six OptiTrack motion capture cameras	Position error	Arduino based mobile robot with 2DOF robotic arm

(Continued)

Table 7. (Continued)

S. No	References	Applications	Algorithms	Sensing Device	Performance Metrics	Manipulator used
3	S. Lee et al. (2019); Chen et al. (2018)	Random objects picking task	Region proposal Network (RPN); RPN with super pixel and CNN	NA	Average precision; Recall; Processing time	Surgical robotic arm
	Sarikaya et al. (2017); Miroslav et al. (2019)		Faster R-CNN	Gripper with camera	Intersection of Union (IoU); Position MSE	KUKA Tiwa (10 DOF), Zymark 3-DOF Robot, Baxter
	Joniak and Muszynski (2017); Dung et al. (2019)		Histogram of Oriented Gradients (HOG) approach			Two HOG wheel mobile robot
4	Li W et al. (2019); Xiao et al. (2017); Minniti et al. (2019)	Obstacle avoidance	Multi constrained Model Predictive control(MMPC)	Four kinetic cameras;	Distance calculation between robot and people;	Gantry robotic manipulator, mobile robot, 4 DOF manipulator Assistive Robotic Manipulator (ARM)
	Cuevas et al. (2020); Benzouai et al. (2016)		Type 2 fuzzy neural network	Proximity sensor; inductive sensor	Linear speed; angular speed; tracking error	Mobile robot, PUMA 560
	Zhou et al. (2018); Wei et al (2019)		CNN	2D LiDAR; Intel Realsense; IMU; camera	Recognition rate; tracking error	Industrial Manipulator, mobile robot
5	Iriondo et al. (2019); Jixin et al. (2019)		Deep Reinforcement Learning (DRL); Deterministic Policy Gradient (DDPG) and Proximal Policy Optimisation (PPO)	NA	Mean rewards; success rate; standard deviation	Traxxas obstacle avoidance car,
	Xu Zhou et al (2019)		Deep Recurrent NN (DRNN)	NA		
	Wang et al. (2018)	Auto sorting system	Region-Based CNN (R-FCN)	NA	Stability analysis	4 DOF planar manipulator
	Basri et al. (2019)		Fast RCNN	Monocular Camera; Intel RealSense D435	Angle tracking	2DOF Industrial Manipulator
					Classification accuracy; confidence value; running time	Industrial Manipulator, Mobile Robot

(Continued)

Table 7. (Continued)

S. No	References	Applications	Algorithms	Sensing Device	Performance Metrics	Manipulator used
6	Vijayan et al (2018)	Pose estimation	Multilayer Perceptron Neural Network, (MLPNN); Radial Basis Function Neural Network (RBFNN); Elman Neural Network (ENN).	Logitech C170; Logitech C270	MSE; Condition number; error histogram	ABB IRB 1200 (6 DOF) manipulator,
	Lin et al. (2018)		Convolutional Neural Network (CNN)	Intel Realsense SR300	Accuracy; translational and rotational errors	Self-made 7-Dof manipulator
	Wang Luo et al (2020)		Deep Q Network (DQN)	VMX450 laser scanning system	Precision, recall, angle error	4 wheeled Mobile vehicle
	Gao et al. (2020)	Twin Delayed Deep Deterministic policy gradients (DDPG)	NA		Convergence evaluation with respect to steps	Turtlebot3

Servoing (IBVS) (Palmieri et al., 2012). The PBVS is a model-based technique where features from the image are transformed into corresponding robot position and orientation in cartesian space whereas in the IBVS method the features in the image space are directly compared with reference features omitting the pose estimation module (Lippiello et al., 2007). The position-based VS can be used when there is a need to track only the target object by making sure that the camera is in its field of view whereas image-based techniques are more robust than PBVS even in the presence of the camera and hand-eye calibration errors (Ghasemi et al., 2019). The manipulated variable for the robot navigation control is contours, features, and corners of the images captured by the robot sensors whereas the control variable is the position or velocity, or angular velocity of the robot relative to the target object (Senapati et al., 2014). To perform the Visual Servoing task, there are two main transformations required namely camera-to-world transformation T_w^c which is obtained from camera calibration of pin-hole camera model or fisheye camera model based on the sensor shape. The second transformation is camera-to-base T_b^c or camera-to-end effector T_e^c which is the homogeneous transformation matrix obtained from hand-eye calibration. This matrix is the combination of translational and rotation vectors of the camera with world coordinates. Since the Visual Servoing task involves the movement of joints in world coordinates, Jacobian matrix J is used to compute the relationship between end effector velocities and joint velocities. The detailed concepts and derivation of Visual Servoing equations can be referred to in (Corke., 1993). There are various estimation techniques like particle filter (Wang Sun et al., 2019), Least Square estimation (Shademan et al., 2010) that are used to compute robust and efficient Jacobian matrix for Visual Servoing applications.

2.4.1. Position-based visual servoing (PBVS)

There are three basic modules of the PBVS system namely Pose Estimation module, Feature Extraction module involving image processing techniques, and Control module in which conventional or intelligent controllers are designed (Banlue et al., 2014) as shown in Figure 11. From the calibrated vision sensor, using camera intrinsic parameters and the geometric 3D model of the object, the pose of the robot end-effector is estimated from the transformation matrix T_E^0 . The features f obtained from the transformation is used to estimate the pose P_a which is a function of the position of the end effector $x(t)$ and orientation θ . The orientation is calculated from inverse kinematics or any heuristic algorithms which calculate the position of the target object $x(t)$ with respect to the camera frame. Finally, the joint controller is designed in such a way that error e between the reference pose P_d and the actual pose P_a becomes zero. Finally, PBVS control calculates linear and angular velocity v_e and ω_e . These values are used to obtain the inverse Jacobian matrix to calculate the joint velocities needed for the robot to move. The advantage of PBVS is that the desired tasks can be formulated in cartesian space which is straightforward to compute the error (Patil et al., 2017). Peng et al. (2020) have done a comparative analysis between PBVS and IBVS for industrial assembly in terms of accuracy and speed. It was implemented in the ABB robot and found that PBVS had faster execution than IBVS and also tested the same with multiple cameras. Similarly, Sharma et al. (2020) implemented PBVS for a mobile robot to estimate the camera to robot base transformation using the gradient descent method. Shi Chen et al (2019) used PBVS for robot pose estimation to avoid obstacles. The authors used the repulsive torque method to find the inverse kinematics of redundant manipulators and found them to be fast, flexible, and accurate.

2.4.2. Image-Based visual servoing (IBVS)

The IBVS approach as illustrated in Figure 12 extracts the features f from the transformation matrix T_E^0 without the need for pose estimation (Keshmiri et al., 2016). In both configurations, the camera can be either fixed in a particular position (eye-to-hand) or placed in the robot end effector (eye-in-hand) configuration. The Image Jacobian Matrix or interaction matrix is calculated which provides transformation mapping between camera velocity V and image feature velocity (Keshmiri et al., 2014). The difference between extracted features f and actual features f^* is calculated to drive the error to zero. In IBVS, there is no need for a pose estimation module and the control law

Figure 11. Closed-loop feedback structure of position-based visual servoing.

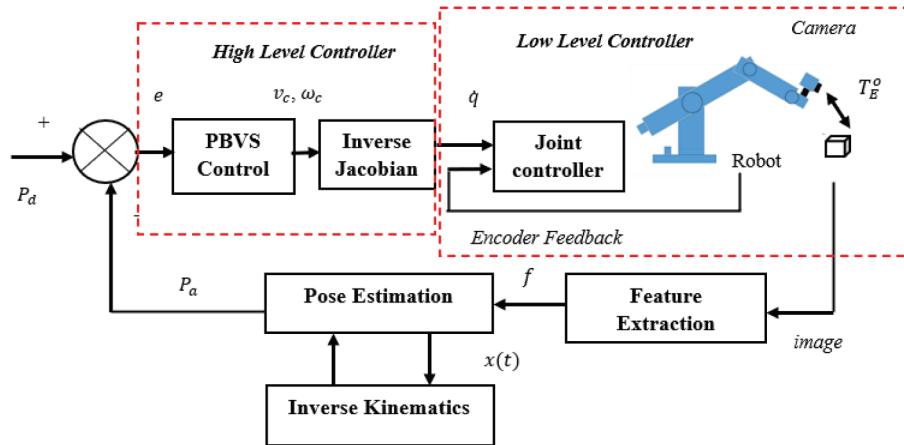
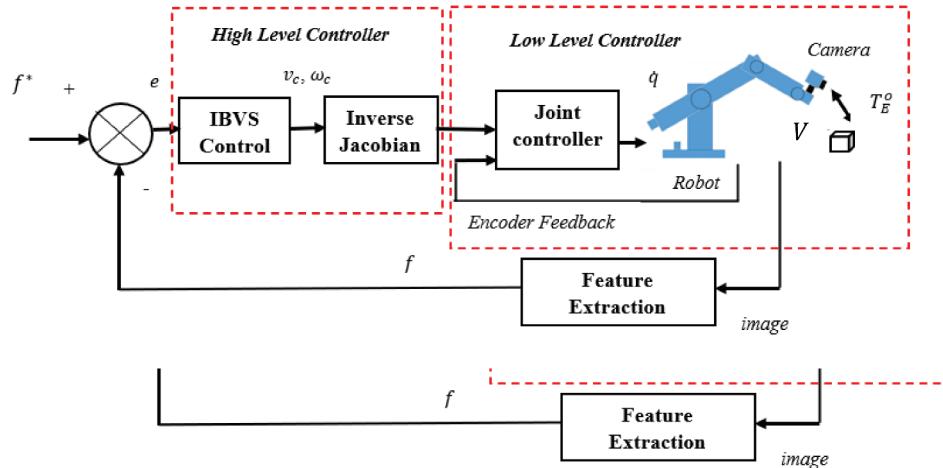


Figure 12. Closed-loop feedback structure of image-based visual servoing.

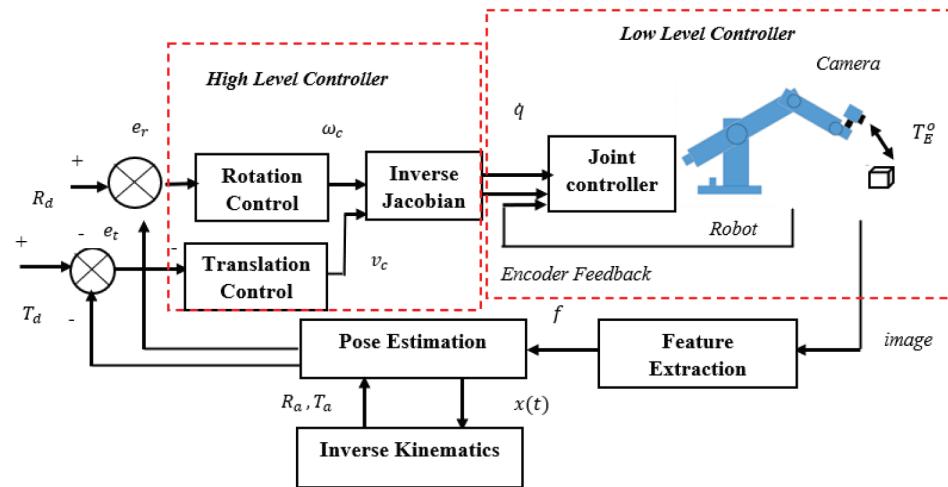


computes velocities of the robot from linear v_c and angular velocity and ω_c of the camera by computing Image Jacobian matrix J_i (Cervera et al., 2003). Though IBVS is robust against errors, there exists some difficulty in IBVS if the camera is disturbed or it is not in a desirable position. If the joints have large rotational errors, IBVS tends to fail. Image singularity, camera retreat, and local minima problem are the other causes of IBVS failure (Nematollahi et al., 2009; Chaumette et al., 1998; Chaumette et al., 2008). Lee et al (2021) proposed dynamic IBVS of wheeled mobile robots using fuzzy integral sliding mode controller to track the feature points in the image plane. By implementing Lyapunov stability analysis transient response was improved. Haviland et al (2020) used IBVS for grasping since other techniques fail when the sensor is close to the object. Using this technique author was able to guide the robot to the object-relative grasp pose using camera RGB information.

2.4.3. 2.5D visual servoing

The pros and cons of IBVS and PBVS which were discussed in the above sections have certain flaws, so researchers combined both IBVS and PBVS into a hybrid approach called 2.5D Visual Servoing (VS) which potentially outperforms these basic schemes (Z. He et al., 2018; Malis et al., 1999) which is shown in Figure 13. One of the most interesting aspects of this control law is that it manages to decouple the rotation and translation control. In this type of VS, the controller is designed such

Figure 13. Closed-loop feedback structure of 2.5d visual servoing.



that rotation error e_r between R_a and R_d as well as a translation error e_t between T_a and T_d becomes zero. The decoupled linear velocity v_c and angular velocity ω_c are fed as inputs to the interaction matrix which gives the joint velocity q as the output. This is then sent to the robot joints to move the end effector in a particular pose. Also, 2.5D Servoing has better stability than IBVS and PBVS, and some of the major disadvantages like the existence of local minima in IBVS, large steady-state errors are removed in PBVS (Gans et al., 2003). W. He et al. (2019) proposed 2.5D Visual Servoing to eliminate higher-order image moments so he implemented this technique for grasping texture less planar parts. The authors used cross-correlation analysis to estimate the position of planar parts and they found that it has stable motion control with 3D trajectory. Q. Li et al. (2020) proposed hybrid VS for harvesting robots to track the unknown fruit motion which is taken as the feature in this work. The authors used an adaptive tracking controller since the motion was dynamic and also performed stability analysis using Lyapunov.

The general idea behind Visual Servoing, its types, and related works has been discussed in depth. The common difference between various methods is consolidated based on the literature survey and it is described in Table 8.

Table 8. Comparison of different visual servoing techniques

S. No	Parameters	PBVS	IBVS	Hybrid
1	Error	Sensitive to calibration errors	Robust to calibration errors	Highly robust
2	Depth	3D trajectory planning can be easily estimated	Depth must be separately estimated	No need of 3D geometric models
3	Input	Position and orientation	Features	Pose and Features
4	Planning	3D Servoing possible	Only 2D Servoing	2D and 3D Servoing
5	Singularity	Low	High	Low
6	Applications	Intelligent navigation	Intelligent transportation	Humanoid robot guidance

2.4.4. Visual servoing based on multiple view geometry

To improve the robustness of Visual Servoing and to generate an optimal trajectory, multiple view geometry techniques can be implemented where feedback signals are processed without prior knowledge about the coordinates of the target. Also, to remove the classical Visual Servoing failure caused by high illumination and calibration errors, and measurement errors, this technique can be effectively used in robot motion control. The different geometrical methods include Homography matrix-based approach, Epipolar geometry, Trifocal-tensor, and Essential matrix-based approach (Li Li et al, 2021).

2.4.4.1. Visual servoing based on homography matrix. The purpose of the homography matrix is to generate an ideal trajectory in image space using projective transformation where 3D coordinates are converted into corresponding 2D coordinates. The scaled Euclidean reconstruction approach is used to calculate the rotation and the translation motion of the camera path. Z. Gong et al. (2020) proposed uncalibrated VS to calculate the shortest path using projective homography which is used to remove large errors. It was tested for Universal Robot 5 where the camera intrinsic and extrinsic are unknown. Wang Zhang et al (2019) used homography for improving robot grasping tasks where the author used grasp images to train Grasp Quality Network (GQN) and the success rate was 78% in real-world testing. Y. Qiu et al. (2017) implemented homography-based VS for tracking control of mobile robots where initially video recorded trajectory is taken as input features and calculated projective geometry of those feature points to perform visual feedback control. To reduce the time cost of image processing and to avoid non-unique solutions, Lopez-Nicolas et al. (2010) used the homography matrix in the control law directly instead of SVD for pose estimation. The main advantages of using homography in visual servoing are it is the simplest algorithm and saves computational time. The limitations of this approach are Z coordinates are assumed zero hence it works for the applications involving planar objects. Next, this method cannot be used if the camera moves out of view hence there is a need to fix the camera where it always sees the camera. But again, it is not practically possible especially with mobile robots where the field of view is dynamic (Dey et al, 2012).

2.4.4.2. Visual servoing based on epipolar geometry. This technique can be useful when two monocular cameras or stereo camera is used to capture the same scene then we use epipolar geometry to calculate the relation between the two views. When the optical centers of each camera are projected onto the other camera's image plane epipolar points are formed (Prasad et al, 2018). Becerra et al (2011) proposed epipolar geometry for dynamic pose estimation of differential drive mobile robot. The pose was further predicted using discrete Extended Kalman Filter (EKF) and it showed better performance than the Essential matrix-based approach. Chesi et al. (2006) used epipolar geometry to estimate the position of apparent contours in unstructured 3D space so that using VS robot can navigate efficiently. Mutib et al (2014) used ANN-based VS to learn the mapping between mobile drifts and visual kinematics and used multi-view epipolar geometry to compensate for the rotational errors. The main advantage of epipolar geometry is that they are easy to implement and provide the relationship between image planes without the knowledge of 3D structure. The only drawback is that they are more susceptible to noise.

2.4.4.3. Visual servoing based on trifocal tensors. In general, the trifocal tensor is a 3x3x3 array that calculates projective geometric relationship in all three dimensions which depends only on the pose and not on scene structure. They are more robust to outliers. Zhang et al proposed a trifocal tensor for Visual Servoing implemented for 6 DOF manipulators. The author built a trifocal tensor between current and reference views to establish a geometric relationship. Finally, a Visual controller with adaptive law was developed to compensate for the unknown distance, and stability was analysed using Lyapunov-based techniques (Zhang et al, 2019). Aranda et al. (2017) developed a 1D tensor model to estimate the angle between omnidirectional images captured at different locations. This technique provided a wide field of view with precise angles. K. K. Lee et al. (2017) used trifocal tensors to estimate the 3D motion of the camera and used two Bayesian filters for point and line measurements. B. Li et al. (2013) develop a 2D trifocal tensor-based visual stabilization controller whereas Chen Jia et al. (2017) designed a trifocal-tensor-based adaptive controller to track the desired trajectory with key video frames as input. The main

advantage is that it is computationally less intensive, but the disadvantage of this method is that no closed-form solution is possible with nonlinear constraints.

The overall summary of Visual Servoing and its different approaches were discussed in depth with advantages and limitations. The main implication of this section is that the execution of robot tasks becomes simpler and efficient when visual information is combined with control theory. Using different image processing techniques like homography matrix, epipolar geometry, Visual Servoing can be optimized. Also, with the help of machine learning and deep learning techniques, it becomes easier to handle large image datasets, planning and control under complex environments can also be achieved.

3. High-Level intelligent controllers for visual servoing

In Visual Servoing, the controller is used to calculate the error e between the reference pose or image feature and actual pose or feature to zero where the visual feedback loop is very effective in increasing the dexterity and flexibility of the robot. Over the past few decades, lots of improvements have been made in the field of control systems applied to robotics because only path planning with the low-level controller is not sufficient. To control the external disturbances like the vibration of motors, jerking, and other environmental factors, optimization, and learning-based approaches are very effective in set-point tracking and stabilization control. The previous literature discussed different conventional approaches but in recent years many visual control schemes are based on optimization for better tracking and motion control of robots. Though conventional controllers like PID are commonly used due to their robustness and simplicity, in robotics the model is non-linear, and to use conventional controllers the model must be linearized. This becomes time-consuming for the complex models. Hence there is a need for optimization techniques that can be combined with conventional controllers for better performance. Further many hybrid approaches using learning-based approaches are being explored. This section explains different works based on optimization approaches and learning-based approaches to design high-level controllers for vision-based control.

3.1. Optimization approaches for visual servoing

The optimization of robot parameters is needed to perform accurate navigation avoiding obstacles. Classical VS techniques cannot handle the constraints effectively and perform unsatisfactorily. Hence VS is combined with optimization algorithms so that large displacements or rotations are handled safely. By taking into account various constraints in robot dynamics and environment, the robot must be able to take its own decisions thereby increasing efficiency, reliability, and stability and reducing the consumption of power and vibrations. The common parameters considered are minimizing velocity, acceleration, and jerk, minimizing the distance between goal and robot, minimizing time, and much more. The optimization problem can be a single or multi-objective problem, Linear or Non-linear. In the multi-objective optimization (MOO) problem, these multiple objectives are treated independently between them giving a set of Pareto optimal solutions which is termed the Pareto set. The common steps in optimization are defining the objective function; selecting the set of variables to maximize or minimize; specifying a set of constraints; choose optimization algorithms like Genetic Algorithm, Least-Squares, Linear Quadratic Regulator, Newton Raphson iteratively until stopping criteria is achieved. There are a few practical constraints in Visual Servoing namely field of view constraint, dynamic saturation constraint, and kinematic saturation constraint (Adrian et al, 2017). For Visual Servoing, enough features must be extracted with respect to the reference image to reduce the error but due to limited FOV of the camera or in case of moving the camera the above criteria get affected. To solve this issue hybrid controllers (Gans et al, 2007) and path planning approaches can be used (Chesi, 2009). Fang et al. (2012) and Freda et al (2007) installed a camera on the pan-tilt platform to keep the target as close as possible to the center of the image. This configuration captures enough features with the target image so that the error signals of the high-level controller become zero. Jia and Liu (2015) removed the redundant FOV by applying homography to system states. Finally, a switching controller is designed to track the shortest path. Jing Xin et al. (2021) proposed improved tracking learning detection (TLD) algorithm which extracts the features when the camera is in FOV or predicts the features when it is away from the FOV. It was experimentally tested in MOTOMAN 6 DOF manipulator and the authors found that it effectively solved the servoing failure due to FOV.

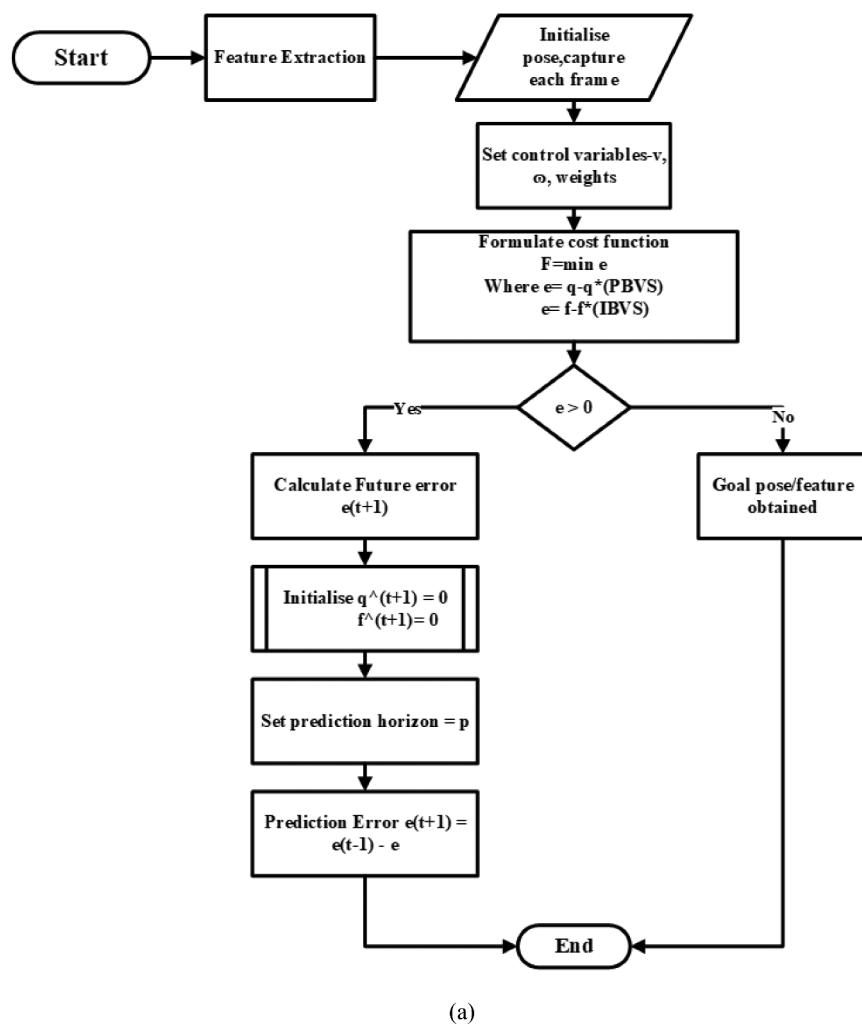
The saturation constraints commonly occur in servo motors when the input signals exceed the maximum limit thereby causing the controller to saturate and degrading the performance of Visual Servoing. To resolve this issue time-varying state feedback control (Jiang et al., 2001), vector field orientation feedback (Michalek et al., 2010), and filters control can be used (X. Chen et al., 2014). Y. Chen et al. (2019) implemented Model Predictive Control (MPC) as a tracking controller with constraints and unknown dynamics to calculate the optimal value of the input signal. Further, the virtual control torque method based on the error state of the dynamic model is regulated to eliminate the effect of saturation and nonlinearity. The servo controllers designed by Wang et al. (2017), (2021)) considered saturation constraint for the design of MPC-based VS. Panah et al (2021) proposed optimized hybrid VS for 7 DOF robot arms to calculate the shorter camera path by eliminating convergence issues, non-optimized trajectories, and singularities. The author used a neuro-fuzzy approach to calculate the pseudoinverse Jacobian matrix and found it to remove singularities making it robust against image noises. Ke et al (2018) implemented a stabilization control strategy for differential drive mobile robots using robust MPC along with an ancillary state feedback system. The author used linear variable inequality-based primal-dual neural network (LVI-PDNN) for gain scheduling and MPC to generate optimized trajectory. J Li et al. (2019) used structured light-based VS to optimize the welding pose. It uses the IBVS method to extract the weld seam center as features and control law was formulated to optimize the orientation and position of the welding torch. The authors found that features were matched, and camera velocities were zero in presence of noises making it highly robust. Y Zhang et al. (2017) used Recurrent Neural Network (RNN) for kinematic control of puma 560 robotic manipulator. The author uses constrained optimization problems with joint and velocity limits and found it to be efficient. Z Qiu et al. (2019) designed MPC for both eye-in-hand and eye-to-hand configurations considering visibility and torque constraints. The authors also designed a sliding mode observer to estimate joint velocity and found that IBVS with MPC gave a satisfactory performance with minimum pixel error. The major drawback of any optimization algorithm is that if the number of control variables is more, the time taken to find the optimal values increases thereby increasing the computational effort. Also, there is a lot of possibilities where the solutions can reach local optimum thereby failing to attain global optimum.

3.2. Learning based approaches for visual servoing

Using artificial intelligence, a set of rules and models were created that allows the robot systems to sense and act like humans in the real world (Peters et al., 2011). The learning approaches for Visual Servoing increase the intelligent ability of robots and a high level and common tasks like human-robot interaction, adapting sudden environmental changes can be done efficiently. Furthermore, the learning approaches are classified into supervised, unsupervised, and reinforcement learning. In supervised learning techniques, both input and targets are known whereas in unsupervised learning only part of the input is known and the target to achieve is unknown. Finally, Reinforcement learning (RL) is used to make decisions sequentially based on rewards obtained from acting on the environment. It can be used for tasks like trajectory optimization, motion planning, dynamic path planning, and Visual Servoing. The effective learning algorithms for Visual Servoing are Deep learning algorithms, RL, Deep RL, and much more. Pandya et al. (2019) proposed a Convolution Neural Network (CNN) based Visual Servoing strategy for vehicle inspection. Shi et al. (2020) implemented a decoupled Visual Servoing with a fuzzy Q-learning module are designed for adjusting mixture parameters for efficient control of wheeled mobile robots. Sampedro et al. (2018) used deep reinforcement learning-based Visual Servoing in object following tasks. Pedersen et al. (2020) proposed DRL for grasping tasks where he initially uses cycleGAN for transferring the raw images into simulated ones and uses DRL for gripper pose estimation and achieved a success rate of 83%. Bateux et al. (2017) proposed deep learning-based VS for the 6DOF robot for estimating the relative pose between two images in the same scene. This method was highly robust even under strong illumination and occlusions and it was able to achieve error less than 1 mm. Castelli et al. (2017) proposed machine learning-based VS that uses Gaussian Mixture Model (GMM) for industrial path following tasks. Saxena et al. (2017) used CNN for end-to-end VS in which knowledge of camera parameters and scene knowledge is unknown. The data used for training CNN was synchronized camera poses and the author tested in UAV where he claims that using CNN there is no need for feature extraction or tracking. MengKang et al. (2020) proposed IBVS based on extreme ML and Q learning to solve the problem of complex modeling. Here ML is used to calculate pseudo inverse interaction matrix to avoid singularities

and uses Q learning to adaptively adjust the gains of servo motors. It was tested in 6 DOF robots with eye-in-hand configuration. Ribeiro et al. (2021) uses CNN for grasping tasks. The authors used Cornell grasping dataset to train CNN to predict the position, orientation, and opening of the robot gripper. Jin et al. (2021) proposed policy-based DRL for IBVS of mobile robots to resolve the problems of feature loss and improve servo efficiency. The flowchart of optimization-based controllers and learning-based controllers has been developed by referring to the literature (Calli B et al., 2018; Bateux et al., 2018; J Li et al., 2019; L Shi et al., 2021) and they are illustrated in Figure 14(a and b) respectively. Though the learning-based approaches can solve complex problems easily there are few limitations. The first and foremost drawback is that it requires lots of data for training to provide better results. Next, it requires powerful processing like GPUs which is not cost-effective. They do not learn interactively and incrementally in real-time.

Figure 14. (a) flowchart for visual servoing using optimization approach **(b)** flowchart for visual servoing using learning-based approaches.



(a)

Figure 14. (a) flowchart for visual servoing using optimization approach (b) flowchart for visual servoing using learning-based approaches.

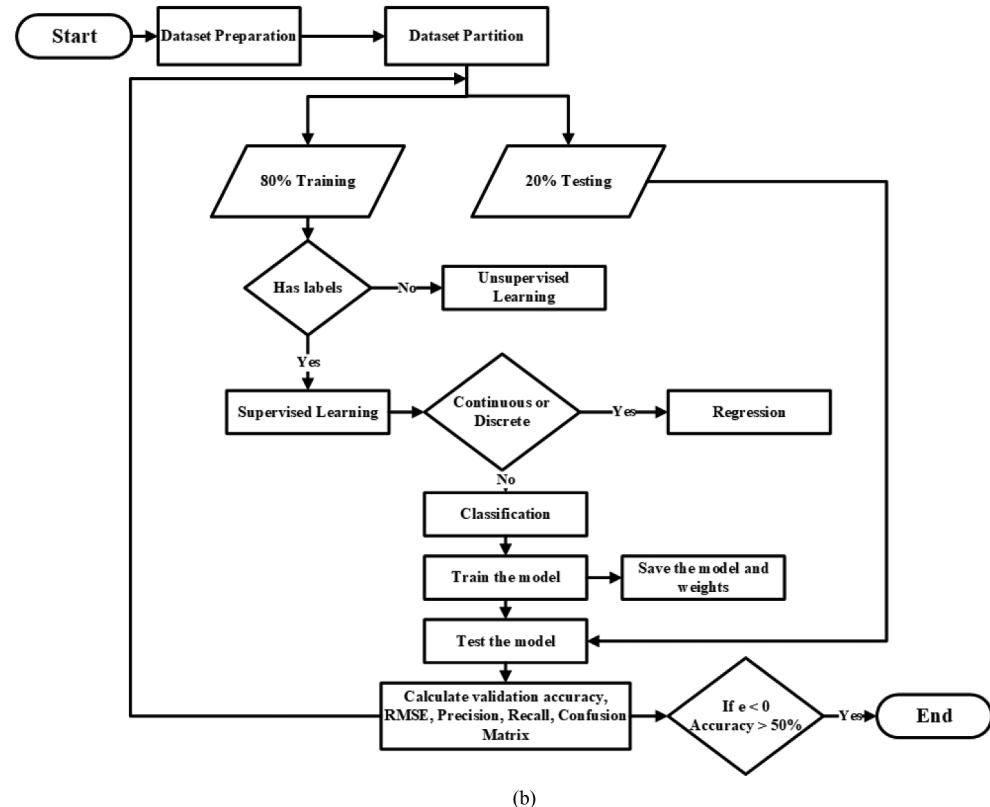


Figure 15. (a) year wise graphical analysis of visual servoing papers published (b) journal database of papers reviewed in this work.

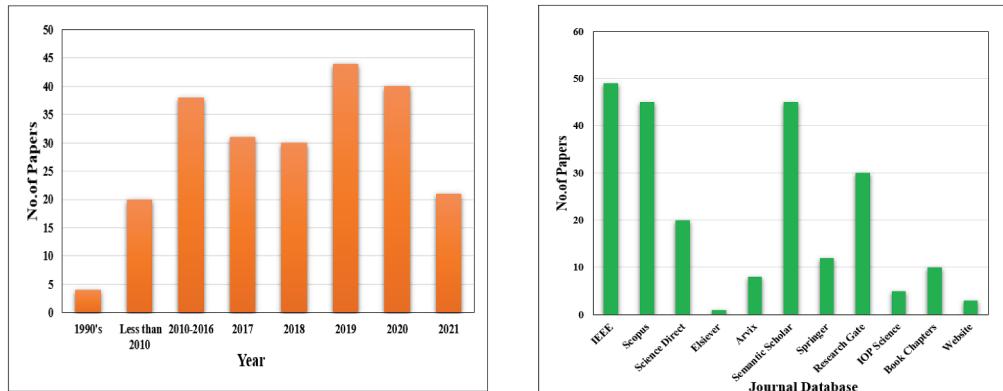
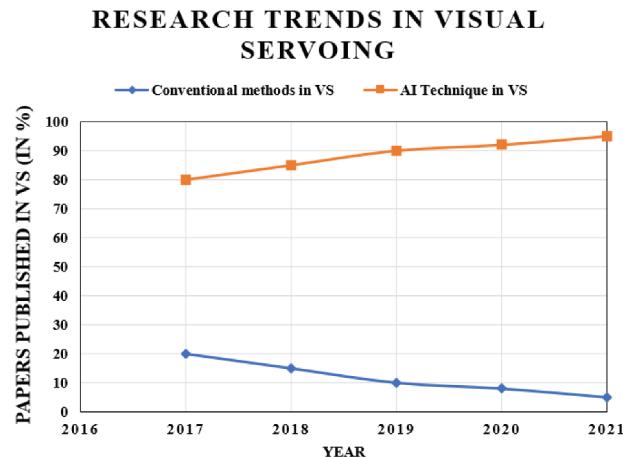


Figure 16. Graphical analysis of visual servoing works between 2017–2021.



4. Literature review analysis in vision-based planning and control

Over the decade, many research papers have been published for vision-based path planning tasks which use both conventional as well as intelligent techniques. This paper reviewed in detail the steps involved in vision-based autonomous systems as a whole dealing with sensors, calibration methodology, path planning, and Visual Servoing. Many papers were reviewed in the area of vision-based path planning and visual feedback controllers using optimization and learning-based approaches. It was found that among papers reviewed there is a huge scope in learning-based approaches particularly in the area of deep learning and deep reinforcement learning for both path planning and Visual Servoing. Most of the papers taken for this review analysis are from well-reputed journals taken from databases namely Scopus, Web of Science, Arvix, and much more and they are recent papers with a duration from 2017–2021. The number of papers investigated for this paperwork is illustrated in the form of graphical analysis based on year wise and journals database referred for this paperwork. From Figures 15(a and b), it can be inferred that more papers were reviewed between 2017–2021. It can also be noted that most of the papers chosen for review analysis were from IEEE and Scopus databases. The reason for choosing such papers is that they gave better insights into the Visual Servoing concept and its various methodologies. It also gave in-depth information about different state-of-art approaches followed in vision-based path planning which helped to formulate an overall review about vision-based autonomous navigation and control. The entire Scopus database contains over 200,000 AI-based papers that are getting published every year where the quality of paper is high. So, the need for AI in the field of robotics is increasing which motivated to formulate this overall review which will help the readers in understanding the concepts of path planning and Visual Servoing. Based on the papers reviewed analysis has been made based on the

Figure 17. (a) Comparison of conventional vs AI approaches between 2000–2012 (b) comparison of conventional vs AI approaches between 2012–2021.

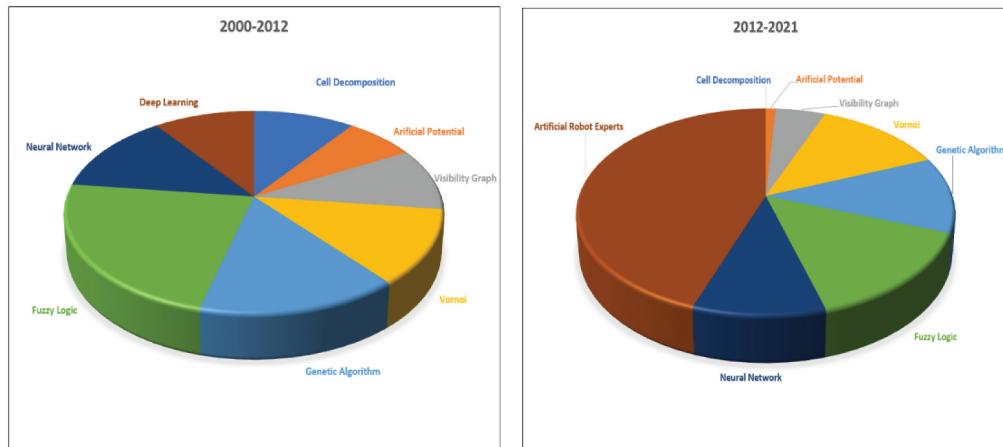
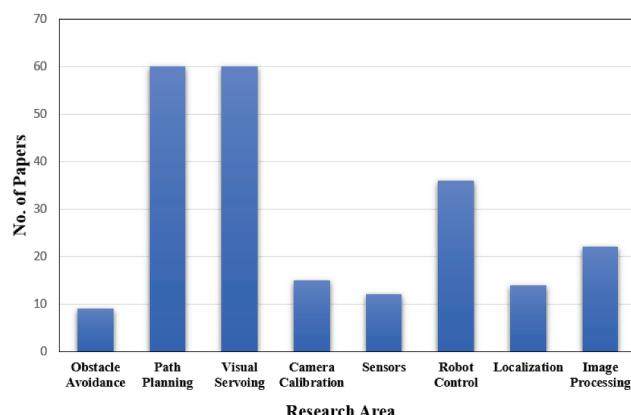


Figure 18. Graphical illustrations of different modules of vision guided robotic systems covered.



number of papers that used conventional and AI approaches in path planning as well as in Visual Servoing. The detailed analysis of research in the field of VS has been plotted year wise in Figure 16.

From the above graph, it can be seen that 80% of research works are carried out in AI-based techniques in the year 2017 which drastically increased to 98% in the year 2021. In the present scenario, the need for robotics in Robot Process Automation is increasing, the research work is carried out based on conventional approach is decreasing because AI can make the robot think and behave like humans, and productivity and quality increases. Also, with developments in processing technology like GPUs, AI algorithm training, and large database handling have become simpler. Furthermore, many simpler tools to perform Visual Servoing in various programming languages have been developed (Riberio et al,2021). Future research work is mainly concentrating on building robots for social interaction with humans and in improving the intelligent ability of robots, the usage of artificial intelligence algorithms is expected to increase diminishing the usage of conventional algorithms. Furthermore, recent research is focussing mainly on unsupervised learning strategies like Dynamic Programming, Intelligent Markov Decision process, Deep Reinforcement Learning under environment. These algorithms tend to make the robot more robust under uncertain conditions making learning better than supervised learning algorithms. Also, three-dimensional Visual Servoing involving point cloud analysis is being researched making computer vision tasks more realistic.

Figure 17 a and b illustrate the trend of conventional and heuristic path planning algorithms for Visual Servoing from the year 2000–2012 and 2012–2021 respectively. During the period from 2000–2012, most of the planning of robots were done using conventional algorithms and deep learning was least explored between the year 1980–2012 which saw a drastic increase from 10% to 40% after 2012. Among them, Voronoi graph-based algorithm was very popular because it was efficient for irregular obstacle shapes. Deep learning was researched more after 2012 after the introduction of transfer learning. From 2012–2021 researchers made a paradigm shift from conventional algorithms to AI-based techniques for path planning because the robot was made to navigate autonomously without hitting the obstacles thereby increasing the intelligence of the robot. It also circumvents the need for modelling the robot kinematics and dynamics using complex mathematical equations because when the number of DOF increases modelling becomes time-consuming. The total contribution of conventional approaches is just 19% whereas intelligent approaches are dominated by about 81% focussing mainly on Artificial Robot Experts.

The recent research shows that artificial intelligence algorithms are dominating in almost all industries since they outperform better for larger datasets. In total, 227 papers have been reviewed to give an overall idea about how the research varied from the conventional approach to AI approaches. This section gives the reader the basic idea behind techniques that were used for the past 20 years in the field of Visual Servoing and vision-based path planning. It can be noticed that before 2012 more conventional

Table 9. Challenges existing in vision-guided robots

S. No	Area	Challenges
1	Dynamic Image reconstruction	Robust 3D reconstruction where both camera and objects are freely moving
2	Video object segmentation (VOS)	Segmenting multiple frames in the video
3	3D hand shape and pose estimation	Manual mesh generation for hand is still challenging
4	Object detection	Need a robust algorithm for detection of large datasets with heavy occlusions and class ambiguities
5	Handwritten recognition	Needs robust dataset to detect different handwriting curves, shapes, and styles
6	Machine vision systems	Need for plug and play systems
7	Quality Inspection	Need for better 3D computer vision tools that could detect biological variability, spectral variation accurately
8	Robot Grasping	Needs novel gripper design and improve cognitive task issues with advanced devices
9	SLAM in unknown Environment	Terrain based navigation in an unconfined environment with accurate maps
10	Multi-Robot Coordination	Need for distributed coordination systems that are flexible than centralized communication

approaches were used to plan the path and perform motion planning. It was the period where robots were operated in manual mode with 100% human intervention. The researchers opted for the Artificial potential field and graph-based approaches for path planning. Further for the design of controllers, PID was commonly used. But though these techniques gave accurate results, it was not able to achieve autonomous behaviour. Also, these algorithms cannot handle complex, dynamic environments to give accurate and mimic realistic situations. This led researchers to shift to learning-based approaches which could increase the intelligent ability of robots. Hence after 2012, with the development of low-cost cameras and processing capabilities, AI-based approaches became popular. To handle dynamic realistic situations Fuzzy logic, ANN, and deep learning were used to perform path planning for Visual Servoing based applications. Using these algorithms robots will also be able to plan the path even in an unknown environment. In recent times, learning-based algorithms combined with optimization algorithms like MPC, GA, Kalman filter techniques are used to predict the future optimal trajectory of the robot efficiently. In further years more research will happen in the field of Reinforcement Learning with deep learning which makes the robot capable of learning from mistakes without human intervention. The number of papers reviewed for each section discussed in this paper is also illustrated in Figure 18.

5. Research gaps, challenges and future scope of vision-based tasks

The usage of robots in both home and industry is accelerating for the past few decades. Based on the statistics released by the International Federation of Robotics, the ratio of robots for every 10,000 human employees, is drastically increasing at a rate of 5% to 9% (Jung et al,2020). But there are still lots of challenges and improvements needed for both industrial and personal robots which researchers are constantly working on it to change the future of manufacturing or for home automation. Among the field of robotics, due to the emergence of low-cost 2D and 3D cameras computer vision has gained popularity for solving problems like tasks Image classification, object detection, image segmentation, path

planning, trajectory generation, Visual Servoing, environment mapping, and so on (Yue et al,2021; Dulari et al,2021; Iuliia et al,2021; Vuuren et al,2020). This section explains the existing challenges in various industrial sectors as well as in computer vision-based tasks from the survey done and a few challenges are listed in Table 9.

The future direction of robotics in Visual Servoing is combining deep learning and reinforcement learning for better feedback control. Also, to make the scene representation more realistic analysis is shifting from two-dimensional to three-dimensional using 3D cameras. Next is combining robots with the Internet of Things(IoT) where robots can be benefitted from cloud services and global connectivity. The only challenges are interoperability, safety, and self-adaptability and further details can be found in (Ray et al,2016; Razafimandimby et al., 2018). There is a need to improve on coordination between humans and robots because human intervention can reduce measurement errors and many complex tasks still require humans for coordination. So, care must be taken to develop a proper architecture for safe and efficient communication between them (Talebpour et al,2019). Furthermore, performance evaluations for Visual Servoing are required to be developed effectively so that robots can be controlled in a better way. For such cases, robustness analysis can be done for sensor failure cases and sensor communication by performing prediction analysis using MPC or using Markov Decision Process to restrict the movement of robots under failure conditions and take decisions accordingly (Verma et al,2021). The optimization algorithms make the system computationally effective with less power consumption which can be used for actions like task allocation, motion planning, and decision making. Other factors which need to be addressed while developing autonomous robots are collision, congestion, and deadlock. Finally, optimal techniques can be combined with 3D image processing and image reconstruction for better real-world understanding.

6. Conclusion

In recent years, there is a huge gap between technology and industrial demands in which the need for fully automated robots is increasing. To remove this gap, industries are shifting towards the implementation of robots in intelligent manufacturing or any other social activities. To satisfy the industrial needs, vision-guided robots are made which are capable of moving accurately and precisely based on the visual feedback received. This in turn enables greater flexibility and higher productivity in a diverse set of manufacturing tasks. The key to successful and efficient vision-based robots can be implemented with better system design which in turn reduces the cost and time-consuming process. Though better technologies have emerged, still there are many issues in vision-based tasks due to illumination and occlusion. So, researchers have been finding many solutions to eliminate this issue. Many works and review papers on different methodologies and their improvements were discussed. But there were very limited works that concentrates on understanding the basic concepts of vision-based robots. Hence this paper aims to provide a detailed survey on different modules of vision-based robots which will be an eye-opener for the readers. The stages which were covered in the discussion was namely perception, localization, path planning, and Visual Servoing. The various stages have been analyzed logically and technically to give clear insights on various techniques and basics involved. The paper also includes designing high-level controllers needed for Visual Servoing using optimization and learning-based approaches to emphasize more on recent approaches as many works on conventional controllers have already been addressed. In each stage, various state-of-art approaches followed in recent years were discussed. The comparison of different camera calibration methodologies which is responsible for better data acquisition has been elaborated. Different path planning algorithms and conventional VS schemes were discussed, and optimization of VS based on multi-view geometry was also covered. The Graphical analysis has also been made year-wise making it convenient for the readers to understand the improvements evolved in the area of Visual Servoing and path planning. Also, the journal database referred for this paperwork is illustrated using a graph to show the quality of papers referred for this work. Finally, the challenges and future direction of Visual Servoing that still exist in vision-guided robots have been discussed. It can be concluded that vision sensors are a good cost-effective solution for vision-based navigation of any robot either industrial manipulators or mobile robots. The overall vision-based robot planning can be used for obstacle detection for safe human-robot interaction,3D environmental mapping, and SLAM-based path planning. From the review, it can be inferred that as the industries are shifting

towards AI and smarter technologies, vision-based robots have more scope where more robust algorithms will be built that can do multiple tasks like human beings and also safely interact with humans. Many technologies involving reinforcement learning are emerging which makes the robot behave and take decisions like a human. The future scope of review works includes a detailed review on behavioral cloning of robots, 3D Visual Servoing, deep reinforcement learning, and SLAM-based techniques.

Acknowledgements

The authors are immensely grateful to the authorities of Birla Institute of Science and Technology Pilani, Dubai campus for their support throughout this research work.

Author details

Abhilasha Singh¹

E-mail: p20180906@dubai.bits-pilani.ac.in

ORCID ID: <http://orcid.org/0000-0003-1270-6249>

V. Kalaichelvi²

ORCID ID: <http://orcid.org/0000-0002-9144-6846>

R. Karthikeyan²

E-mail: rkarthikeyan@dubai.bits-pilani.ac.in

ORCID ID: <http://orcid.org/0000-0001-7401-7698>

¹ Department of Electrical and Electronics Engineering, Birla Institute of Technology and Science Pilani, Dubai Campus, Dubai, UAE, P.O.Box. 345 055.

² Department of Mechanical Engineering, Birla Institute of Technology and Science Pilani, Dubai Campus, Dubai, UAE, P.O.Box. 345 055.

Cover Image

Source: Author.

Disclosure statement

No potential conflict of interest was reported by the author(s). <https://www.ijser.org/onlineResearchPaperViewer.aspx?A-Review-on-Wireless-Embedded-System-for-Vision-Guided-Robot-Arm-for-Object-Sorting.pdf> IEEE

Funding

The authors did not receive any funding for this work.

Citation information

Cite this article as: A survey on vision guided robotic systems with intelligent control strategies for autonomous tasks, Abhilasha Singh, V. Kalaichelvi & R. Karthikeyan, *Cogent Engineering* (2022), 9: 2050020.

References

- Adrian, O. D., Serban, O. A., & Niculae, M. (2017). Some proper methods for optimization in robotics. *International Journal of Modeling and Optimization*, 7(4), 188–193. <https://doi.org/10.7763/IJMO.2017.V7.582>
- Aguiar, A. S., Dos Santos, F. N., Cunha, J. B., Sobreira, H., & Sousa, A. J. (2020). Localization and mapping for robots in agriculture and forestry: A survey. *Robotics*, 9(4), 1–23. <https://doi.org/10.3390/robotics9040097>
- Al Mutib, K., Mattar, E., Alsulaiman, M., Ramdane, H., & ALDean, M. (2014). Neural network vision-based visual servoing and navigation for KSU-IMR mobile robot using epipolar geometry. *The International Journal of Soft Computing and Software Engineering*, March 1-2 (pp. 1–8). March 1-2 3(March 1-2. <https://doi.org/10.7321/jcse.v3.n3.136>.
- Al-Forati, I. S., & Rashid, A. (2018). Design and implementation of an indoor robot localization system using minimum bounded circle algorithm. In 2019 8th International Conference on Modeling Simulation and Applied Optimization (ICMSAO), Apr 15-17, (pp. 1–6). IEEE, Manama, Bahrain.
- Al-Jarra, R., Al-Jarra, M., & Roth, H. (2018). A novel edge detection algorithm for mobile robot path planning. *Journal of Robotics*, 2018, 1–12. <https://doi.org/10.1155/2018/1969834>
- Ali, I., Durmush, A., Suominen, O., Yli-Hietanen, J., Peltonen, S., Collin, J., & Gotchev, A. (2020). FinnForest dataset: A forest landscape for visual SLAM. *Robotics and Autonomous Systems*, 132, 1–13. <https://doi.org/10.1016/j.robot.2020.103610>
- Ali, H., Gong, D., Wang, M., & Dai, X. (2020). Path planning of mobile robot with improved ant colony algorithm and MDP to produce smooth trajectory in grid-based environment. *Frontiers in Neurorobotics*, 14(44), 1–13. <https://doi.org/10.3389/fnbot.2020.00044>
- Almeida, J. S., Marinho, L. B., Souza, J. M., Assis, E. A., & Reboucas Filho, P. P. (2018). Localization system for autonomous mobile robots using machine learning methods and omnidirectional sonar. *IEEE Latin America Transactions*, 16(2), 368–374. <https://doi.org/10.1109/LA.2018.8327388>
- Alom, M. Z., Taha, T. M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M. S., Hasan, M., Van Essen, B. C., Awwal, A. A. S., & Asari, V. K. (2019). A state-of-the-art survey on deep learning theory and architectures. *Electronics*, 8(3), 1–67. <https://doi.org/10.3390/electronics8030292>
- Angani, A., Lee, J. W., Talluri, T., Lee, J. Y., & Shin, K. J. (2020). Human and Robotic Fish Interaction Controlled Using Hand Gesture Image Processing. *Sensors and Materials*, 32(10), 3479–3490. <https://doi.org/10.18494/SAM.2020.2925>
- Aranda, M., Lopez-Nicolas, G., & Sagües, C. (2017). Angle-based navigation using the 1D trifocal tensor. In Aranda, M., Lopez-Nicolas, G., & Sagües, C. (Eds.), *Control of Multiple Robots Using Vision Sensors* (vol. 34, no. 1, pp. 19–51). Springer. https://doi.org/10.1007/978-3-319-57828-6_2
- Arhin, S., Manandhar, B., & Baba-Adam, H. (2020). Predicting travel times of bus transit in Washington, DC using artificial neural networks. *Civil Engineering Journal*, 6(11), 2245–2261. <https://doi.org/10.28991/cej-2020-03091615>
- Asvadi, A., Garrote, L., Premebida, C., Peixoto, P., & Nunes, U. J. (2017). Multimodal vehicle detection: Fusing 3D LIDAR and color camera data. *Pattern Recognition Letters*, 115(1), 9–20. <https://doi.org/10.1016/j.patrec.2017.09.038>
- Banlue, T., Sooraksa, P., & Noppanakepong, S. (2014). A practical position-based visual servo design and implementation for automated fault insertion test. *international Journal of Control, Automation and Systems*, 12(5), 1090–1101. <https://doi.org/10.1007/s12555-013-0128-3>
- Basri, H., Syarif, I., Sukardhoto, S., & Falah, M. F. (2019). Intelligent system for automatic classification of fruit defect using faster region-based convolutional neural network (Faster R-Cnn). *Jurnal Ilmiah Kursor*, 10(1), 1–11. <https://doi.org/10.28961/kursor.v10i1.187>
- Bateux, Q., Marchand, E., Leitner, J., Chaumette, F., & Corke, P. (2017). Visual servoing from deep neural networks. *arXiv preprint, arXiv:1705.08940*. 1–6. <https://doi.org/10.48550/arXiv.1705.08940>
- Bateux, Q., Marchand, E., Leitner, J., Chaumette, F., & Corke, P. (2018). Training deep neural networks for

- visual servoing. In *2018 IEEE international conference on robotics and automation (ICRA)*, May 21–25, (pp. 3307–3314). IEEE: Brisbane, Australia. <https://doi.org/10.1109/ICRA.2018.8461068>.
- Bebis, G., Egbert, D., & Shah, M. (2003). Review of computer vision education. *IEEE Transactions on Education*, 46(1), 2–21. <https://doi.org/10.1109/TE.2002.808280>
- Becerra, H. M., & Sagues, C. (2011). Dynamic pose-estimation from the epipolar geometry for Visual Servoing of mobile robots, In *2011 IEEE International Conference on Robotics and Automation*, 2011, May 9–13, (pp. 417–422). IEEE: Shanghai, China. <https://doi.org/10.1109/ICRA.2011.5979737>.
- Benzaoui, M., Chekireb, H., Tadjine, M., & Boulkroune, A. (2016). Trajectory tracking with obstacle avoidance of redundant manipulator based on fuzzy inference systems. *Neurocomputing*, 196, 23–30. <https://doi.org/10.1016/j.neucom.2016.02.037>
- Bore, N., Jensfelt, P., & Folkesson, J. (2018). Multiple object detection, tracking and long-term dynamics learning in large 3D maps. Arxiv. 1–13. <https://arxiv.org/abs/1801.09292>.
- Bovik, A. C. (2009d). *The Essential Guide to Image Processing*. Academic Press.
- Byambasuren, B. E., Baasanjav, T., Myagmarjav, T., & Baatar, B. (2020). Application of image processing and industrial robot arm for quality assurance process of production, In *2020 IEEE Region 10 Symposium (TENSYMP)*, Jun 5–7, (pp. 526–530), IEEE, Dhaka, Bangladesh.
- Calli, B., & Dollar, A. M. (2018). Robust precision manipulation with simple process models using visual servoing techniques with disturbance rejection. *IEEE Transactions on Automation Science and Engineering*, 16(1), 406–419. <https://doi.org/10.1109/TASE.2018.2819661>
- Campbell, S., Mahony, N. O., Carvalho, A., Kralikova, L., Riordan, D., & Walsh, J. (2020). Path Planning Techniques for Mobile Robots A Review, In *2020 6th International Conference on Mechatronics and Robotics Engineering (ICMRE)*, Feb 12–15, (pp. 12–16). IEEE: Barcelona, Spain, <https://doi.org/10.1109/ICMRE49073.2020.9065187>.
- Castelli, F., Michieletto, S., Ghidoni, S., & Pagello, E. (2017). A machine learning-based visual servoing approach for fast robot control in industrial setting. *International Journal of Advanced Robotic Systems*, 14(6), 1–10. <https://doi.org/10.1177/1729881417738884>
- Cervera, E., Del Pobil, A. P., Berry, F., & Martinet, P. (2003). Improving image-based visual servoing with three-dimensional features. *The International Journal of Robotics Research*, 22(10–11), 821–839. <https://doi.org/10.1177/027836490302210003>
- Chaumette, F. 1998. Potential problems of stability and convergence in image-based and position-based visual servoing. In David J. Kriegman, Gregory D. Hage, A. Stephen Morse (Eds.), *The confluence of vision and control* (Vol. 237 66–78). Springer.
- Chaumette, F., & Hutchinson, S. (2008). Visual Servoing and visual tracking. In (Ed.), *Springer Handbook of Robotics* (pp. 563–583). Springer.
- Chen, X., Jia, Y., & Matsuno, F. (2014). Tracking control of nonholonomic mobile robots with velocity and acceleration constraints. In *2014 American Control Conference*, Jun 4–6, (pp. 880–884). IEEE: Portland, Oregon.
- Chen, J., Jia, B., & Zhang, K. (2017). Trifocal tensor-based adaptive visual trajectory tracking control of mobile robots. *IEEE Transactions on Cybernetics*, 47(11), 3784–3798. <https://doi.org/10.1109/TCYB.2016.2582210>
- Chen, Y., Li, Z., Kong, H., & Ke, F. (2019). Model predictive tracking control of nonholonomic mobile robots with coupled input constraints and unknown dynamics. *IEEE Transactions on Industrial Informatics*, 15(6), 3196–3205. <https://doi.org/10.1109/TII.2018.2874182>
- Chen, B., & Pan, B. (2020). Camera calibration using synthetic random speckle pattern and digital image correlation. *Optics and Lasers in Engineering*, 126, 1–9. <https://doi.org/10.1016/j.optlaseng.2019.105919>
- Chen, B. X., Sahdev, R., & Tsotsos, J. K. (2017). Integrating Stereo Vision with a CNN Tracker for a Person-Following Robot. In M. Liu, H. Chen, & M. Vincze (Eds.), *ICVS 2017. Lecture Notes in Computer Science* (Vol. 10528, pp. 300–313). Springer: Computer Vision Systems.
- Chen, J., Wang, Z., & Li, H. (2018). Real-time object segmentation based on convolutional neural network with saliency optimization for picking. *Journal of Systems Engineering and Electronics*, 29(6), 1300–1307. <https://doi.org/10.21629/JSEE.2018.06.17>
- Chesi, G. (2009). Visual Servoing path planning via homogeneous forms and LMI optimizations. *IEEE Transactions on Robotics*, 25(2), 281–291. <https://doi.org/10.1109/TRO.2009.2014131>
- Chesi, G., Mariottini, G. L., Prattichizzo, D., & Vicino, A. (2006). Epipole-based Visual Servoing for mobile robots. *Advanced Robotics*, 20(2), 55–280. <https://doi.org/10.1163/156855306775525820>
- Chopde, S., Patil, M., Shaikh, A., Chavhan, B., & Deshmukh, M. (2017). Developments in the computer vision system, focusing on its applications in quality inspection of fruits and vegetables-A review. *Agricultural Reviews*, 38(2), 94–102. <https://doi.org/10.18805/ag.v38i02.7940>
- Corke, P. I. (1993). Visual control of robot manipulators—a review. *Visual Servoing: Real-Time Control of Robot Manipulators Based on Visual Sensory Feedback*, 7, 31. https://doi.org/10.1142/9789814503709_0001
- Cruz, Y. J., Rivas, M., Quiza, R., Beruvides, G., & Haber, R. E. (2020). Computer vision system for welding inspection of liquefied petroleum gas pressure vessels based on combined digital image processing and deep learning techniques. *Sensors*, 16, 1–13. <https://doi.org/10.3390/s20164505>
- Cuevas, F., Castillo, O., & Cortes, P. (2020). Towards a control strategy based on type-2 fuzzy logic for an autonomous mobile robot. In O. Castillo & P. Melin (Eds.), *Hybrid Intelligent Systems in Control, Pattern Recognition and Medicine. Studies in Computational Intelligence* (Vol. 827, pp. 301–314). Springer.
- Cui, X., Lu, C., & Wang, J. (2020). 3D semantic map construction using improved ORB-SLAM2 for mobile robot in edge computing environment. *IEEE Access*, 8, 67179–67191. <https://doi.org/10.1109/ACCESS.2020.2983488>
- D', E. G., & Gasbarro, D. D. (2017). Review of techniques for 2D camera calibration suitable for industrial vision systems. *Journal of Physics: Conference Series*, 841(1), 1–7. <https://doi.org/10.1088/1742-6596/841/1/012030>
- Dai, X., Long, S., Zhang, Z., & Gong, D. (2019). Mobile robot path planning based on ant colony algorithm with a* heuristic method. *Frontiers in Neurorobotics*, 13(15), 131–139. <https://doi.org/10.3389/fnbot.2019.00015>
- Dan, Z., & Bin, W. (2017). A review on model reference adaptive control of robotic manipulators. *Annual Reviews in Control*, 44, 1–11. <https://doi.org/10.1016/j.arcontrol.2017.02.002>
- Dewi, T., Risma, P., Oktarina, Y., & Muslimin, S. (2018). Visual servoing design and control for agriculture robot; a review. In *2018 International Conference on Electrical Engineering and Computer Science (ICECOS)*,

- Oct. 24-25, (pp. 57–62). IEEE: Jakarta, Indonesia.
<https://doi.org/10.1109/ICECOS.2018.8605209>.
- Dey, S., Reilly, V., Saleemi, I., & Shah, M. (2012). Detection of independently moving objects in non-planar scenes via multi-frame monocular epipolar constraint. A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, & C. Schmid (Eds.), *Computer Vision – ECCV 2012. ECCV 2012. Lecture Notes in Computer Science* (Vol. 7576, pp. 860–873). Springer: VerlagBerlin, Heidelberg. https://doi.org/10.1007/978-3-642-33715-4_62
- Dolgov, D., Thrun, S., Montemerlo, M., & Diebel, J. (2008). Practical search techniques in path planning for autonomous driving. *The Ann Arbor Journal*, 1001 (48105), 18–80. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.475.6856>
- Doulamis, A., Doulamis, N., Protopapadakis, E., & Voudelimos, A. (2018) Combined convolutional neural networks and fuzzy spectral clustering for real time crack detection in tunnels, In *2018 25th IEEE International Conference on Image Processing (ICIP)*, Oct 7-10, (pp. 4153–4157). IEEE: Athens, Greece.
- Dung, T. D., Hossain, D., Kaneko, S. I., & Capi, G. (2019). Multifeature image indexing for robot localization in textureless environments. *Robotics*, 8(2), 1–11. <https://doi.org/10.3390/robotics8020037>
- Fang, Y., Liu, X., & Zhang, X. (2012). Adaptive active visual servoing of nonholonomic mobile robots. *IEEE Transactions on Industrial Electronics*, 59(1), 486–497. <https://doi.org/10.1109/TIE.2011.2143380>
- Fareh, R., Baziyad, M., Rahman, M. H., Rabie, T., & Bettayeb, M. (2020). Investigating reduced path planning strategy for differential wheeled mobile robot. *Robotica*, 38(2), 235–255. <https://doi.org/10.1017/S0263574719000572>
- Faria, D. R., Vieira, M., Premebida, C., & Nunes, U. (2015). Probabilistic human daily activity recognition towards robot-assisted living. In *Proceedings of the IEEE ROMAN’15*, Aug 31 – Sep 4, (pp. 582–587). IEEE: Kobe, Japan. <https://doi.org/10.1109/ROMAN.2015.7333644>.
- Fauadi, M. H., Akmal, S., Ali, M. M., Anuar, N. I., Ramli, S., Noor, A. Z., & Awang, N. (2018). Intelligent vision-based navigation system for mobile robot: A technological review. *Periodicals of Engineering and Natural Sciences (PEN)*, 6(2), 47–57. <https://doi.org/10.2153/pen.v6i2.174>
- Fernandes, R., Premebida, C., Peixoto, P., Wolf, D., & Nunes, U. (2014). Road detection using high resolution LIDAR. In *IEEE Vehicle Power and Propulsion Conference*, Oct. 27-30, (pp.1-6). Coimbra, Portugal.
- Ferreira, M., Costa, P., Rocha, L., & Moreira, A. P. (2014). Stereo-based real-time 6-DoF work tool tracking for robot programming by demonstration. *International Journal of Advanced Manufacturing and Technology*, 85(1-4), 1–13. <https://doi.org/10.1007/s00170-014-6026-x>
- Fong, T., Nourbakhsh, I., & Dautenhahn, K. (2003). A survey of socially interactive robots. *Robotics and Autonomous Systems*, 42(3-4), 143–166. [https://doi.org/10.1016/S0921-8890\(02\)00372-X](https://doi.org/10.1016/S0921-8890(02)00372-X)
- Freida, L., & Oriolo, G. (2007). Vision-based interception of a moving target with a nonholonomic mobile robot. *Robotics and Autonomous Systems*, 55(6), 419–432. <https://doi.org/10.1016/j.robot.2007.02.001>
- Gans, N. R., & Hutchinson, S. A. (2007). Stable Visual Servoing through hybrid switched-system control. *IEEE Transactions on Robotics*, 23(3), 530–540. <https://doi.org/10.1109/TRO.2007.895067>
- Gans, N. R., Hutchinson, S., & Corke, P. I. (2003). Performance tests for visual servo control systems, with application to partitioned approaches to visual servo control. *The International Journal of Robotics Research*, 22(10–11), 955–984. <https://doi.org/10.1177/027836490302210011>
- Gao, J., Ye, W., Guo, J., & Li, Z. (2020). Deep reinforcement learning for indoor mobile robot path planning. *Sensors*, 20(19), 1–15. <https://doi.org/10.3390/s20195493>
- Garibotto, G., Murrieri, P., Capra, A., De Muro, S., Petillo, U., Flammini, F., Esposito, M., Pragliola, C., Di Leo, G., Lengu, R., & Mazzino, N. (2013). White paper on industrial applications of computer vision and pattern recognition. In *International conference on image analysis and processing*, Sep 11-13, (pp. 721–730). Springer: Naples, Italy.
- Garrido, S., Moreno, L., Blanco, D., & Jurewicz, P. (2011). Path planning for mobile robot navigation using Voronoi diagram and fast marching. *International Journal of Robotics and Automation*, 2(1), 42–64. <https://doi.org/10.1109/IROS.2006.282649>
- Gasparetto, A., Boscaroli, P., Lanzutti, A., & Vidoni, R. (2015). Path planning and trajectory planning algorithms: A general overview. *Mechanisms and Machine Science*, 29, 3–27. https://doi.org/10.1007/978-3-319-14705-1_1
- Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley.
- Gong, Z., Tao, B., Qiu, C., Yin, Z., & Ding, H. (2020). Trajectory planning with shortest path for modified uncalibrated visual servoing based on projective homography. *IEEE Transactions on Automation Science and Engineering*, 17(2), 1076–1083. <https://doi.org/10.1109/TASE.2019.2954598>
- Gong, L., Yu, X., & Wang, J. (2021). Curve-Localizability-SVM active localization research for mobile robots in outdoor environments. *Applied Science*, 11(10), 1–14. <https://doi.org/10.3390/app11104362>
- Gul, F., Mir, I., Abualigah, L., Sumari, P., & Forestiero, A. (2021). A consolidated review of path planning and optimization techniques: technical perspectives and future directions. *Electronics*, 10(18), 1–38. <https://doi.org/10.3390/electronics10182250>
- Gunen, M. A., Besdok, E., Civicioglu, P., & Atasever, U. H. (2020). Camera calibration by using weighted differential evolution algorithm: A comparative study with ABC, PSO, COBIDE, DE, CS, GWO, TLBO, MVMO, FOA, LSHADE, ZHANG, and BOUQUET. *Neural Computing and Applications*, 32(23), 17681–17701. <https://doi.org/10.1007/s00521-020-04944-1>
- Guo, Y., Wang, X., Xu, Q., Liu, S., Liu, S., & Han, J. (2020). Weather impact on passenger flow of rail transit lines. *Civil Engineering Journal*, 6(2), 276–284. <https://doi.org/10.28991/cej-2020-03091470>
- Halme, R.-J., Lanz, M., Kamaraainen, J., Pieters, R., Latokartano, J., & Hietanen, A. (2018). Review of vision-based safety systems for human-robot collaboration. *Procedia CIRP*, 72(51), 111–116. <https://doi.org/10.1016/j.procir.2018.03.043>
- Haralick, R. M., & Shapiro, L. G. (1993). *Computer and Robot Vision*. Prentice Hall.
- Hartley, R. I. (1997). Self-calibration of stationary camera. *International Journal of Computer Vision*, 22(1), 5–23. <https://doi.org/10.1023/A:1007957826135>
- Haviland, J., Dayoub, F., & Corke, P. (2020). Control of the final-phase of closed-loop visual grasping using image-based visual servoing. *arXiv preprint arXiv:2001.05650*.
- He, W., Heng, W., Jiahai, H., Bin, Z., & Long, Q. (2019). Smooth point-to-point trajectory planning for industrial robots with kinematical constraints based on high-order polynomial curve. *Mechanism and*

- Machine Theory*, 139, 284–293. <https://doi.org/10.1016/j.mechmachtheory.2019.05.002>
- He, Z., Wu, C., Zhang, S., & Zhao, X. (2018). Moment-based 2.5-D visual servoing for textureless planar part grasping. *IEEE Transactions on Industrial Electronics*, 66(10), 7821–7830. <https://doi.org/10.1109/TIE.2018.2886783>
- Hu, J., Niu, H., Carrasco, J., Lennox, B., & Arvin, F. (2020). Voronoi-based multi-robot autonomous exploration in unknown environments via deep reinforcement learning. *IEEE Transactions on Vehicular Technology*, 69(12), 14413–14423. <https://doi.org/10.1109/TVT.2020.3034800>
- Hua, J., & Zeng, L. (2021). Hand-Eye calibration algorithm based on an optimized neural network. *Actuators*, 10 (4), 1–11. <https://doi.org/10.3390/act10040085>
- Huang, W., Liu, H., & Wan, W. (2020). An online initialization and self-calibration method for stereo visual-inertial odometry. *IEEE Transactions on Robotics*, 36(4), 1153–1170. <https://doi.org/10.1109/TRO.2019.2959161>
- Huang, Y., & Su, J. (2019). Visual servoing of nonholonomic mobile robots: a review and a novel perspective. *IEEE Access*, 7, 134968–134977. <https://doi.org/10.1109/ACCESS.2019.2941962>
- Iriondo, A., Lazcano, E., Susperregi, L., Urain, J., Fernandez, A., & Molina, J. (2019). Pick and place operations in logistics using a mobile manipulator controlled with deep reinforcement learning. *Applied Sciences*, 9(348), 1–19. <https://doi.org/10.3390/app9020348>
- Jackson, W. (2016). *Digital Video Editing Fundamentals*. Apress.
- Jansen-van Vuuren, R. D., Shahnewaz, A., & Pandey, A. K. (2020). Image and signal sensors for computing and machine vision: Developments to meet future needs. In Oleg Sergiyenko, Wendy Flores-Fuentes, Paolo Mercorelli (Eds.), *Machine Vision and Navigation* (pp. 3–32). Springer.
- Jia, B., & Liu, S. (2015). Switched visual servo control of nonholonomic mobile robots with field-of-view constraints based on homography. *Control Theory Technology*, 13(4), 311–320. <https://doi.org/10.1007/s11768-015-4068-8>
- Jiang, Z., Lefeber, E., & Nijmeijer, H. (2001). Saturated stabilization and tracking of a nonholonomic mobile robot. *System and Control Letters*, 42(5), 327–332. [https://doi.org/10.1016/S0167-6911\(00\)00104-3](https://doi.org/10.1016/S0167-6911(00)00104-3)
- Jiang, J., & Ma, Y. (2020). Path planning strategies to optimize accuracy, quality, build time and material use in additive manufacturing: A Review. *Micromachines*, 11(7), 1–18. <https://doi.org/10.3390/mi11070633>
- Jiexin, X., Zhenzhou, S., Yue, L., Yong, G., & Jindong, T. (2019). Deep reinforcement learning with optimized reward functions for robotic trajectory planning. *IEEE Access*, 7, 105669–105679. <https://doi.org/10.1109/ACCESS.2019.2932257>
- Jin, Z., Wu, J., Liu, A., Zhang, W. A., & Yu, L. (2021). Policy-based deep reinforcement learning for visual servoing control of mobile robots with visibility constraints. *IEEE Transactions on Industrial Electronics*, 69(2), 1898–1908. <https://doi.org/10.1109/TIE.2021.3057005>
- Joniak, P., & Muszyński, R. (2017). Path following for two HOG wheels mobile robot. *Journal of Automation Mobile Robotics and Intelligent Systems*, 11(2), 75–81. https://doi.org/10.14313/JAMRIS_2-2017/19
- Jung, J. H., & Lim, D. G. (2020). Industrial robots, employment growth, and labor cost: A simultaneous equation analysis. *Technological Forecasting and Social Change*, 159, 1–11. <https://doi.org/10.1016/j.techfore.2020.120202>
- Kakani, V., Nguyen, V. H., Kumar, B. P., Kim, H., & Pasupuleti, V. R. (2020). A critical review on computer vision and artificial intelligence in food industry. *Journal of Agriculture and Food Research*, 2, 1–12. <https://doi.org/10.1016/j.jafr.2020.100033>
- Kanellakis, C., & Nikolakopoulos, G. (2017). Survey on computer vision for UAVs: Current developments and trends. *Journal of Intelligent and Robotic Systems*, 87 (1), 141–168. <https://doi.org/10.1007/s10846-017-0483-z>
- Kang, M., Chen, H., & Dong, J. (2020). Adaptive visual servoing with an uncalibrated camera using extreme learning machine and Q-leaning. *Neurocomputing*, 402, 384–394. <https://doi.org/10.1016/j.neucom.2020.03.049>
- Kazemi, M., Gupta, K., & Mehrandez, M. (2010). Path-planning for visual servoing: A review and issues. *Visual Servoing via Advanced Numerical Methods*, 401, 189–207. https://doi.org/10.1007/978-1-84996-089-2_11
- Ke, F., Li, Z., & Yang, C. (2017). Robust tube-based predictive control for visual servoing of constrained differential-drive mobile robots. *IEEE Transactions on Industrial Electronics*, 65(4), 3437–3446. <https://doi.org/10.1109/TIE.2017.2756595>
- Keshmiri, M., & Xie, W. F. (2016). Image-based Visual Servoing using an optimized trajectory planning technique. *IEEE/ASME Transactions on Mechatronics*, 22(1), 359–370. <https://doi.org/10.1109/TMECH.2016.2602325>
- Keshmiri, M., Xie, W. F., & Mohebbi, A. (2014). Augmented image-based visual servoing of a manipulator using acceleration command. *IEEE Transactions on Industrial Electronics*, 61(10), 5444–5452. <https://doi.org/10.1109/TIE.2014.2300048>
- Lai, C. C., & Su, K. L. (2018). Development of an intelligent mobile robot localization system using Kinect RGB-D mapping and neural network. *Computers and Electrical Engineering*, 7, 620–628. <https://doi.org/10.1016/j.compeleceng.2016.04.018>
- Le, A. V., Kyaw, P. T., Veerajagadheswar, P., Muthugala, M. V., Elara, M. R., Kumar, M., & Nhan, N. H. (2021). Reinforcement learning-based optimal complete water-blasting for autonomous ship hull corrosion cleaning system. *Ocean Engineering*, 220, 1–16. <https://doi.org/10.1016/j.oceaneng.2020.108477>
- Lee, J. Y. (2020). Zero-Shot Calibration of Fisheye Cameras. *arXiv preprint arXiv:2011.14607*.
- Lee, S., & Chwa, D. (2021). Dynamic image-based visual servoing of monocular camera mounted omnidirectional mobile robots considering actuators and target motion via fuzzy integral sliding mode control. *IEEE Transactions on Fuzzy Systems*, 29(7), 2068–2076. <https://doi.org/10.1109/TFUZZ.2020.2985931>
- Lee, S., Kim, H., Ko, D., Lee, M., Jung, B., & Seok, S. (2019). Trajectory tracking of end effector on mobile robot with multiple onboard cameras. In 2019 16th International Conference on Ubiquitous Robots (UR), Jun 24–27, (pp. 212–218). IEEE: Jeju, Korea.
- Lee, D.-M., Yang, S.-H., Lee, S.-R., & Lee, Y.-M. (2004). Development of machine vision system and dimensional analysis of the automobile front-chassis-module. *KSME International Journal*, 18(12), 2209–2215. <https://doi.org/10.1007/BF02990225>
- Lee, K. K., Yu, Y. K., & Wong, K. H. (2017). Recovering camera motion from points and lines in stereo images: A recursive model-less approach using trifocal tensors. In 2017 18th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), Jun 26–28, (pp. 83–90). IEEE: Kanazawa, Japan. <https://doi.org/10.1109/SNPD.2017.8022704>.

- Li, G. C.-H., & Chang, Y.-M. (2019). Automated visual positioning and precision placement of a workpiece using deep learning. *The International Journal of Advanced Manufacturing Technology*, 104(9), 4527–4538. <https://doi.org/10.1007/s00170-019-04293-x>
- Li, J., Chen, Z., Rao, G., & Xu, J. (2019). Structured light-based visual servoing for robotic pipe welding pose optimization. *IEEE Access*, 7, 138327–138340. <https://doi.org/10.1109/ACCESS.2019.2943248>
- Li, B., Fang, Y., & Zhang, X. (2013). 2d trifocal tensor based visual servo regulation of nonholonomic mobile robots. In *32nd IEEE Chinese Control Conference*, Jul 26-28, (pp. 5764–5769). IEEE: Xi'an, China.
- Li, C., Li, B., Wang, R., & Zhang, X. (2021). A survey on visual servoing for wheeled mobile robots. *International Journal of Intelligent Robotics and Applications*, 5(2), 203–218. <https://doi.org/10.1007/s41315-021-00177-0>
- Li, Q., Nevalainen, P., Queraltà, J. P., Heikkonen, J., & Westerlund, T. (2020). Localization in Unstructured Environments: Towards Autonomous Robots in Forests with Delaunay Triangulation. *Remote Sensing*, 12(11), 1–22. <https://doi.org/10.3390/rs12111870>
- Li, T., Qiu, Q., & Zhao, C. (2021). Hybrid Visual Servoing Tracking Control of Uncalibrated Robotic Systems for Dynamic Dwarf Culture Orchards Harvest. In *2021 IEEE International Conference on Development and Learning (ICDL)*, Aug 23-26, (pp. 1–6). IEEE: Beijing, China.
- Li, W., & Xiong, R. (2019). Dynamical Obstacle Avoidance of Task-Constrained Mobile Manipulation Using Model Predictive Control. *IEEE Access*, 7, 88301–88311. <https://doi.org/10.1109/ACCESS.2019.2925428>
- Lin, H. (2014). A Fast and Unified Method to Find a Minimum-Jerk Robot Joint Trajectory Using Particle Swarm Optimization. *Journal of Intelligent and Robotic Systems*, 75(3–4), 379–392. <https://doi.org/10.1007/s10846-013-9982-8>
- Lin, C.-M., Tsai, C.-Y., Lai, Y.-C., Li, S.-A., & Wong, -C.-C. (2018). Visual Object Recognition and Pose Estimation Based on a Deep Semantic Segmentation Network. *IEEE Sensors Journal*, 18(22), 9370–9381. <https://doi.org/10.1109/JSEN.2018.2870957>
- Lippiello, V., Siciliano, B., & Villani, L. (2007). Position-Based Visual Servoing in Industrial Multirobot Cells Using a Hybrid Camera Configuration. *IEEE Transactions on Robotics*, 23(1), 73–86. <https://doi.org/10.1109/TRO.2006.886832>
- Liu, H., Darabi, H., Banerjee, P., & Liu, J. (2007). Survey of wireless indoor positioning techniques and systems. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 37(6), 1067–1080. <https://doi.org/10.1109/TSMCC.2007.905750>
- Long, L., & Dongri, S. (2019). Review of Camera Calibration Algorithms. In Sanjiv K. Bhatia, Shailesh Tiwari, Su Ruidan, Munesh Chandra Trivedi, K. K. Mishra (Eds.), *Advances in Computer Communication and Computational Sciences* (pp. 723–732). Springer.
- López-Nicolás, G., Gans, N. R., Bhattacharya, S., Sagües, C., Guerrero, J. J., & Hutchinson, S. (2010). Homography-based control scheme for mobile robots with non-holonomic and field-of-view constraints. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 40(4), 1115–1127. <https://doi.org/10.1109/TSMCB.2009.2034977>
- Luo, L. P., Yuan, C., Yan, R. J., Yuan, Q., Wu, J., Shin, K. S., & Han, C. S. (2015). Trajectory planning for energy minimization of industry robotic manipulators using the Lagrange interpolation method. *International Journal of Precision Engineering and Manufacturing*, 16(5), 911–917. <https://doi.org/10.1007/s12541-015-0119-9>
- Malamas, Elias, Petrakis, Euripides, Zervakis, Michalis, Petit, Laurent, Legat, Jean-Didier. (2003). A survey on industrial vision systems, applications, and tools, *Image and Vision Computing*, 21, 171–188. [https://doi.org/10.1016/S0893-9682\(02\)00152-X](https://doi.org/10.1016/S0893-9682(02)00152-X)
- Malis, E., Chaumette, F., & Bouvet, S. (1999). 2 1/2 D Visual Servoing. *IEEE Transactions on Robotics and Automation*, 15(2), 238–250. <https://doi.org/10.1109/70.760345>
- Man, L., Yong, D., Xianli, H., & Maolin, Y. (2020). Image positioning and identification method and system for coal and gangue sorting robot. *International Journal of Coal Preparation and Utilization*, 14(11), 1–19. <https://doi.org/10.1080/19392699.2020.1760855>
- Manduchi, R., Castano, A., Talukder, A., & Matthies, L. (2005). Obstacle detection and terrain classification for autonomous off-road navigation. *Autonomous Robots*, 18(1), 81–102. <https://doi.org/10.1023/B:AURO.0000047286.62481.1d>
- Martinez, E. A., Caron, G., Pegard, C., & Alabazares, D. L. (2020). Photometric path planning for vision-based navigation. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 31 May to 31 August, (pp. 9007–9013). IEEE: Paris, France. <https://doi.org/10.1109/ICRA40945.2020>
- Mezouar, Y., & Chaumette, F. (2002). Path planning for robust image-based control. *IEEE Transactions on Robotics and Automation*, 18(4), 534–549. <https://doi.org/10.1109/TRA.2002.802218>
- Michalek, M., & Kozłowski, K. (2010). Vector-field-orientation feedback control method for a differentially driven vehicle. *IEEE Transactions on Control Systems Technology*, 18(1), 45–65. <https://doi.org/10.1109/TCST.2008.2010406>
- Minniti, M. V., Farshidian, F., Grandia, R., & Hutter, M. (2019). Whole-body mpc for a dynamically stable mobile manipulator. *IEEE Robotics and Automation Letters*, 4(4), 3687–3694. <https://doi.org/10.1109/LRA.2019.2927955>
- Miroslav, S., Karel, K., Miroslav, K., Viktor, K., & Libor, P. (2019). Visual Data Simulation for Deep Learning in Robot Manipulation Tasks, (Ed.): MESAS 2018. Springer. 402–411.
- Mittal, R. K., Nagrath, I. J. (2003). *Robotics and Control*. Tata McGraw-Hill.
- Morar, A., Moldoveanu, A., Mocanu, I., Moldoveanu, F., Radoi, I. E., Asavei, V., Grădinaru, A., & Butean, A. (2020). A comprehensive survey of indoor localization methods based on computer vision. *Sensors*, 20(9), 1–36. <https://doi.org/10.3390/s20092641>
- Mouli, C. C., & Raju, K. N. (2013). A review on wireless embedded system for vision guided robot arm for object sorting. *International Journal of Scientific & Engineering Research*, 4(7), 1–6. <https://www.ijser.org/onlineResearchPaperViewer.aspx?A-Review-on-Wireless-EMBEDDED-System-for-Vision-Guided-Robot-Arm-for-Object-Sorting.pdf>
- Mustafa, C. B., & Omur, A. (2020). Performing predefined tasks using the human–robot interaction on speech recognition for an industrial robot. *Engineering Applications of Artificial Intelligence*, 95, 1–13. <https://doi.org/10.1016/j.engappai.2020.103903>
- Nadhir Ab Wahab, M., Nefti-Meziani, S., & Atyabi, A. (2020). A comparative review on mobile robot path planning: Classical or meta-heuristic methods? *Annual Reviews in Control*, 50, 233–252. <https://doi.org/10.1016/j.arcontrol.2020.10.001>
- Nagrath, M. &, & Nagrath, I. J. (2003). *Robotics and Control*. Tata. McGraw-Hill Education.
- Nandini, V., Deepak Vishal, R., Arun Prakash, C., & Aishwarya, S. (2016). A Review on Applications of

- Machine Vision Systems in Industries. *Indian Journal of Science and Technology*, 9(48), 1–5. <https://doi.org/10.17485/ijst/2016/v9i48/108433>
- Nematollahi, E., & Janabi-Sharifi, F. (2009). Generalizations to control laws of image-based visual servoing. *International Journal of Optomechatronics*, 3(3), 167–186. <https://doi.org/10.1080/15599610903144161>
- Neto, A. A., Macharet, D. G., Da.s. Campos, V. C., & Campos, M. F. M. (2009). Adaptive complementary filtering algorithm for mobile robot localization. *Journal of the Brazilian Computer Society*, 15(3), 19–31. <https://doi.org/10.1007/BF03194503>
- Nuzzi, C., Pasinetti, S., Lancini, M., Docchio, F., & Sansoni, G. (2019). Deep learning-based hand gesture recognition for collaborative robots. *IEEE Instrumentation and Measurement Magazine*, 22(2), 44–51. <https://doi.org/10.1109/MIM.2019.8674634>
- Palmieri, G., Palpacelli, M., Battistelli, M., & Callegari, M. (2012). A comparison between position-based and image-based dynamic visual servoings in the control of a translating parallel manipulator. *Journal of Robotics*, 2012, 1–11. <https://doi.org/10.1155/2012/103954>
- Pan, X., Wu, J., Li, Z., Yang, J., Zhang, C., Deng, C., Yi, Z., Gao, Q., Yu, M., Zhang, Z., Liu, L., Chi, F., & Bai, P. (2021). Self-calibration for linear structured light 3D measurement system based on quantum genetic algorithm and feature matching. *Optik*, 225, 1–10. <https://doi.org/10.1016/j.ijleo.2020.165749>
- Pandey, A., Pandey, S., & Parhi, D. R. (2017). Mobile Robot Navigation and Obstacle Avoidance Techniques: A Review. *International Robotics & Automation Journal*, 2 (3), 1–12. <https://doi.org/10.15406/iratj.2017.02.00023>
- Pandya, H., Gaud, A., Kumar, G., & Krishna, K. M. (2019). Instance invariant visual servoing framework for part-aware autonomous vehicle inspection using MAVs. *Journal of Field Robotics*, 36(5), 892–918. <https://doi.org/10.1002/rob.21859>
- Parikh, D. (2020, April 27). Computer Vision Tools and Libraries. The Research Nest . <https://medium.com/the-research-nest/computer-vision-tools-and-libraries-52bb34023bdf>
- Patel, P., & Bhavsar, B. (2021). Object Detection and Identification. *International Journal of Advanced Trends in Computer Science and Engineering*, 10(3), 1611–1618. <https://doi.org/10.30534/ijatcse/2021/181032021>
- Patil, M. (2017). Robot Manipulator Control Using PLC with Position Based and Image Based Algorithm. *International Journal of Swarm Intelligence and Evolutionary Computation*, 6(1), 1–8. <https://doi.org/10.4172/2090-4908.1000154>
- Patle, B. K., Ganesh Babu, L., Pandey, A., Parhi, D. R. K., & Jagadeesh, A. (2019). A review: On path planning strategies for navigation of mobile robot. *Defence Technology*, 15(4), 582–606. <https://doi.org/10.1016/j.dt.2019.04.011>
- Paya, L., Gil, A., & Reinoso, O. (2017). A State-of-the-Art Review on Mapping and Localization of Mobile Robots Using Omnidirectional Vision Sensors. *Journal of Sensors*, 2017, 1–20. <https://doi.org/10.1155/2017/3497650>
- Pedersen, O. M., Misimi, E., & Chaumette, F. (2020). Grasping unknown objects by coupling deep reinforcement learning, generative adversarial networks, and visual servoing. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 31 May–31 Aug, (pp. 5655–5662). IEEE: Paris, France. <https://doi.org/10.1109/ICRA40945.2020.9197196>
- Peng, Y.-C., Jivani, D., Radke, R. J., & Wen, J. (2020). Comparing Position- and Image-Based Visual Servoing for Robotic Assembly of Large Structures. In *2020 IEEE 16th International Conference on Automation Science and Engineering (CASE)*, Aug 20–21, (pp. 1608–1613). IEEE: Hong Kong, China. <https://doi.org/10.1109/CASE48305.2020.9217028>
- Peters, J., Tedrake, R., Roy, N., & Morimoto, J. (2011). Robot learning. In *Encyclopedia of Machine Learning* (pp. 865–869). Springer: Bruno Siciliano, Oussama Khatib.
- Pirahansiah, F., Abdullah, S. N., & Sahran, S. (2015). Camera calibration for multi-modal robot vision based on image quality assessment. In *2015 10th Asian Control Conference (ASCC)*, May 31–Jun 3, (pp. 1–6). IEEE: Kota Kinabalu, Malaysia.
- Prasad V., Das D., & Bhowmick B. (2018). Epipolar Geometry based Learning of Multi-view Depth and Ego-Motion from Monocular Sequences. arXiv: 1812.11922, 1–10. <https://doi.org/10.1145/3293353.3293427>
- Premebida, C., Ambrus, R., & Marton, Z. C. (2018). Intelligent robotic perception systems. In Efren Gorroño Hurtado (Ed.) *Applications of Mobile Robots*. IntechOpen. <https://doi.org/10.5772/intechopen.79742>
- Premebida, C., & Nunes, U. (2013). Fusing LIDAR, camera and semantic information: A context-based approach for pedestrian detection. *The International Journal of Robotics Research*, 32(3), 371–384. <https://doi.org/10.1177/0278364912470012>
- Qingyang, C., Zhenping, S., Daxue, L., Yuqiang, F., & Xiaohui, L. (2012). Local path planning for an unmanned ground vehicle based on SVM. *International Journal of Advanced Robotic Systems*, 9 (6), 1–13. <https://doi.org/10.5772/54130>
- Qiu, Z., Hu, S., & Liang, X. (2019). Model predictive control for uncalibrated and constrained image-based visual servoing without joint velocity measurements. *IEEE Access*, 7, 73540–73554. <https://doi.org/10.1109/ACCESS.2019.2920389>
- Qiu, Y., Li, B., Shi, W., & Chen, Y. (2017). Homography-based visual servo tracking control of wheeled mobile robots with simultaneous depth identification. In *International Conference on Neural Information Processing*, 324–333. Guangzhou, China: Springer. https://doi.org/10.1007/978-3-319-70136-3_35
- Rastegarpanah, A., Aflakian, A., & Stolkin, R. (2021). Optimized hybrid decoupled visual servoing with supervised learning. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, 236(2), 338–354. <https://doi.org/10.1177/0956518211028379>
- Ray, P. P. (2016). Internet of robotic things: Concept, technologies, and challenges. *IEEE Access*, 4, 9489–9500. <https://doi.org/10.1109/ACCESS.2017.2647747>
- Razafimandimby, C., Loscri, V., & Vigni, A. M. (2018). Towards efficient deployment in Internet of Robotic Things. In Raffaele Gravina, Carlos E. Palau, Marco Manso, Antonio Liotta, Giancarlo Fortino (Eds.), *Integration, interconnection, and interoperability of IoT systems* (pp. 21–37). Springer.
- Reis, R., Santos, F. N., & Santos, L. (2019). Forest Robot and Datasets for Biomass Collection. In Manuel F. Silva, Jose Luis Lima, Luis Paulo Reis, Alberto Sanfeliu, Danilo Tardoli (Eds.), *Advances in Intelligent Systems and Computing* (pp. 152–163). Springer Nature.
- Ribeiro, E. G., de Queiroz Mendes, R., & Grassi, V., Jr. (2021). Real-time deep learning approach to visual servo control and grasp detection for autonomous robotic manipulation. *Robotics and Autonomous Systems*, 139, 1–24. <https://doi.org/10.1016/j.robot.2021.103757>

- Richard Hartley, R. G., & Chang, T. (1992). Stereo from uncalibrated cameras. In *Proceedings of Computer Vision and Pattern Recognition*, Jun 15-18, (vol. 92, pp. 761–764). IEEE: Champaign, IL, USA.
- Rigoberto, J.-S., Juan, Z., & Victor, H. D.-R. (2020). Distorted pinhole camera modeling and calibration. *Applied Optics*, 59(36), 11310–11318. <https://doi.org/10.1364/AO.412159>
- Rout, A., Deepak, B. V. L., & Biswal, B. B. (2019). Advances in weld seam tracking techniques for robotic welding: A review. *Robotics and Computer Integrated Manufacturing*, 56, 12–37. <https://doi.org/10.1016/j.rcim.2018.08.003>
- Saarinen, J., Andreasson, H., & Lilienthal, A. J. (2012). Independent Markov chain occupancy grid maps for representation of dynamic environment. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, May 14-19, (pp. 3489–3495). IEEE: Vilamoura-Algarve, Portugal.
- Sampedro, C., Rodriguez-Ramos, A., Gil, I., Mejias, L., & Campoy, P. (2018). Image-based visual servoing controller for multirotor aerial robots using deep reinforcement learning. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Oct 1-5, pp. 979–986, IEEE: Madrid, Spain. <https://doi.org/10.1109/IROS.2018.8594249>
- Santos, L. C., Aguiar, A. S., Santos, F. N., Valente, A., & Petry, M. (2020). Occupancy grid and topological maps extraction from satellite images for path planning in agricultural robots. *Robotics*, 9(4), 1–22. <https://doi.org/10.3390/robotics9040077>
- Santos, L. C., Santos, F. N., Solteiro Pires, E. J., Valente, A., Costa, P., & Magalhaes, S. (2020). Path Planning for ground robots in agriculture: A short review. In *2020 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)*, Apr 15-17, (pp. 61–66). IEEE: Azores, Portugal. <https://doi.org/10.1109/ICARSC49921.2020.9096177>
- Sarabu, A., & Santra, A. K. (2021). Human Action recognition in videos using convolution long short-term memory network with spatio-temporal networks. *Emerging Science Journal*, 5(1), 25–33. <https://doi.org/10.28991/esj-2021-01254>
- Sarapura, A., Roberti, F., Gimenez, J., Patino, D., & Carelli, R. (2018). Adaptive visual servoing control of a manipulator with uncertainties in vision and dynamics. In *2018 Argentine Conference on Automatic Control (AADECA)*, Nov 7-9, (pp. 1–6). Buenos Aires, Argentina. <https://doi.org/10.23919/AADECA44134.2018>
- Sarikaya, D., Corso, J. J., & Guru, K. A. (2017). Detection and localization of robotic tools in robot-assisted surgery videos using deep neural networks for region proposal and detection. *IEEE Transactions on Medical Imaging*, 36(7), 1542–1549. <https://doi.org/10.1109/TMI.2017.2665671>
- Saxena, A., Pandya, H., Kumar, G., Gaud, A., & Krishna, K. M. (2017). Exploring convolutional networks for end-to-end visual servoing. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, 29 May-3 Jun, (pp. 3817–3823). Singapore: IEEE. <https://doi.org/10.1109/ICRA.2017.7989442>
- Senapati, M., Srinivas, J., & Balakrishnan, V. (2014). Visual servoing and motion control of a robotic device in inspection and replacement tasks. *International Journal of Engineering and Technical Research (IJETR)*, 2(4), 37–43. https://www.academia.edu/7725817/Visual_Servoing_and_Motion_control_of_a_robotic_device_in_inspection_and_replacement_tasks
- Shademan, A., Farahmand, A., & Jagersand, M. (2010). Robust Jacobian estimation for uncalibrated Visual Servoing. In *2010 IEEE International Conference on Robotics and Automation*, Nov 7-9, (pp. 5564–5569). Anchorage, AK, USA: IEEE. <https://doi.org/10.1109/ROBOT.2010.5509911>
- Sharma, R., Shukla, S., Behera, L., & Subramanian, V. (2020). Position-based visual servoing of a mobile robot with an automatic extrinsic calibration scheme. *Robotica*, 38(5), 831–844. <https://doi.org/10.1017/S0263574719001115>
- Shi, H., Chen, J., Pan, W., Hwang, K., & Cho, Y. (2019, September). Collision Avoidance for Redundant Robots in Position-Based Visual Servoing. *IEEE Systems Journal*, 13(3), 3479–3489. <https://doi.org/10.1109/JSYST.2018.2865503>
- Shi, L., Copot, C., & Vanlanduit, S. (2021). A Bayesian deep neural network for safe visual servoing in human-robot interaction. *Frontiers in Robotics and AI*, 8, 1–13. <https://doi.org/10.3389/frobt.2021.687031>
- Shi, H., Xu, M., & Hwang, K. S. (2019). A fuzzy adaptive approach to decoupled visual servoing for a wheeled mobile robot. *IEEE Transactions on Fuzzy Systems*, 28 (12), 3229–3243. <https://doi.org/10.1109/TFUZZ.2019.2931219>
- Song, L., Ma, H., Wu, M., Zhou, Z., & Fu, M. (2018). A brief survey of dimension reduction. In *International Conference on Intelligent Science and Big Data Engineering*, Oct 18-20, (pp. 189–200). Lanzhou, China: Springer.
- Su, Q., Yu, W., & Liu, J. (2021). Mobile robot path planning based on improved ant colony algorithm. In *2021 Asia-Pacific Conference on Communications Technology and Computer Science (ACCTCS)*, Jan 22-24, (pp. 220–224). Shenyang, China: IEEE.
- Sunderhauf, N., Dayoub, F., McMahon, S., Talbot, B., Schulz, R., Corke, P., Wyeth, G., Upcroft, B., & Milford, M. (2016). Place categorization and semantic mapping on a mobile robot. In *IEEE International Conference on Robotics and Automation (ICRA)*, May 16-21, (pp. 5729–5736). Stockholm, Sweden: IEEE.
- Talebpour, Z., & Martinoli, A. (2019). Adaptive risk-based replanning for human-aware multi-robot task allocation with local perception. *IEEE Robotics and Automation Letters*, 4(4), 3790–3797. <https://doi.org/10.1109/LRA.2019.2926966>
- Tang, S. H., Khaksar, W., Ismail, N. B., & Ariffin, M. K. (2012). A review on robot motion planning approaches. *Journal of Science and Technology*, 20(1), 15–29. [http://www.pertanika.upm.edu.my/resources/files/Pertanika%20PAPERS%20Vol.%202020%20\(1\)%20Jan.%202012/07%20Pg%2015-29.pdf](http://www.pertanika.upm.edu.my/resources/files/Pertanika%20PAPERS%20Vol.%202020%20(1)%20Jan.%202012/07%20Pg%2015-29.pdf)
- Trigatti, G., Boscaroli, P., Scalera, L., Pillan, D., & Gasparetto, A. (2018). A new path-constrained trajectory planning strategy for spray painting robots - rev.1. *The International Journal of Advanced Manufacturing Technology*, 98(9-12), 2287–2296. <https://doi.org/10.1007/s00170-018-2382-2>
- Triggs, B. (1997). Autocalibration and the absolute quadric. *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Juan, PR, USA, 17-19 June 1997, IEEE. <https://doi.org/10.1109/CVPR.1997.609388> doi:10.1109/CVPR.1997.609388
- Tsai, R. Y. (1987). A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE Journal of Robotics and Automation*, 3(4), 323–344. <https://doi.org/10.1109/JRA.1987.1087109>
- Tzafestas Spyros, G. (2018). Mobile robot control and navigation: A global overview. *Journal of Intelligent*

- and Robotic Systems, 91(1), 35–58. <https://doi.org/10.1007/s10846-018-0805-9>
- Verma, J. K., & Ranga, V. (2021). Multi-robot coordination analysis, taxonomy, challenges and future scope. *Journal of Intelligent and Robotic Systems*, 102(1), 1–36. <https://doi.org/10.1007/s10846-021-01378-2>
- Vijayan, A. T., & Ashok, S. (2018). Comparative study on the performance of neural networks for prediction in assisting Visual Servoing. *Journal of Intelligent & Fuzzy Systems*, 36(1), 675–688. <https://doi.org/10.3233/JIFS-171475>
- Wang, X., Fang, G., Wang, K., Xie, X., Lee, K. H., Ho, T. W. L., Kwok, J. L., W., K., & Kwok, K.-W. (2020). Eye-in-hand visual servoing enhanced with sparse strain measurement for soft continuum robots. *IEEE Robotics and Automation Letters*, 5(2), 2161–2168. <https://doi.org/10.1109/LRA.2020.2969953>
- Wang, B., Liu, Z., Li, Q., & Prorok, A. (2020). Mobile robot path planning in dynamic environments through globally guided reinforcement learning. *IEEE Robotics and Automation Letters*, 5(4), 6932–6939. <https://doi.org/10.1109/LRA.2020.3026638>
- Wang, W., Luo, H., Zheng, Q., Wang, C., & Guo, W. (2020). A deep reinforcement learning framework for vehicle detection and pose estimation in 3D point clouds. In International Conference on Artificial Intelligence and Security, pp. 405–416, Hohhot, China: Springer. https://doi.org/10.1007/978-3-030-57881-7_36
- Wang, F., Sun, F., Zhang, J., Lin, B., & Li, X. (2019). Unscented particle filter for online total image Jacobian matrix estimation in robot visual servoing. *IEEE Access*, 7, 92020–92029. <https://doi.org/10.1109/ACCESS.2019.2927413>
- Wang, T., Wang, W., & Wei, F. (2020, October 25). An overview of control strategy and trajectory planning of visual servoing. In Recent Featured Applications of Artificial Intelligence Methods, LSMS 2020 and ICSEE 2020 Workshops, (pp. 358–370). Shanghai, China: Springer.
- Wang, T., Yao, Y., Chen, Y., Zhang, M., Tao, F., & Snoussi, H. (2018). Auto-Sorting System Toward Smart Factory Based on Deep Learning for Image Segmentation. *IEEE Sensors Journal*, 18(20), 8493–8501. <https://doi.org/10.1109/JSEN.2018.2866943>
- Wang, R., Zhang, X., & Fang, Y. (2021). Visual tracking of mobile robots with both velocity and acceleration saturation constraints. *Mechanical Systems and Signal Processing*, 150, 1–16. <https://doi.org/10.1016/j.ymssp.2020.107274>
- Wang, R., Zhang, X., Fang, Y., & Li, B. (2017). Visual Servoing of mobile robots with input saturation at kinematic level. In 2017 International Conference on Image and Graphics, Sep 13–15, (pp. 432–442). Shanghai, China: Springer Cham.
- Wang, X., Zhang, X., Ren, X., Li, L., Feng, H., He, Y., Chen, H., & Chen, X. (2020). Point cloud 3D parent surface reconstruction and weld seam feature extraction for robotic grinding path planning. *The International Journal for Advanced Manufacturing Technologies*, 107(1), 827–841. <https://doi.org/10.1007/s00170-020-04947-1>
- Wang, A. S., Zhang, W., Troniak, D., Liang, J., & Kroemer, O. (2019). Homography-based deep visual servoing methods for planar grasps. In 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Nov 4–8, (pp. 6570–6577). Macau, China: IEEE. <https://doi.org/10.1109/IROS40897.2019.8968160>
- Wang, X., Zhao, Y., & Yang, F. (2019). Camera calibration method based on Pascal's theorem. *International Journal of Advanced Robotic Systems*, 16(3), 1–10. <https://doi.org/10.1177/1729881419846406>
- Wu, P., Cao, Y., He, Y., & Li, D. (2017). Vision-Based Robot Path Planning with Deep Learning. In M. Liu, H. Chen, & M. Vincze (Eds.), *Computer Vision Systems. ICSV 2017. lecture notes in computer science* (Vol. 10528, pp. 1–11). Springer.
- Xiao, H., Li, Z., Yang, C., Zhang, L., Yuan, P., Ding, L., & Wang, T. (2017). Robust stabilization of a wheeled mobile robot using model predictive control based on neurodynamics optimization. *IEEE Transactions on Industrial Electronics*, 64(1), 505–516. <https://doi.org/10.1109/TIE.2016.2606358>
- Xie, J., Shao, Z., Li, Y., Guan, Y., & Tan, J. (2019). Deep reinforcement learning with optimized reward functions for robotic trajectory planning. *IEEE Access*, 7, 105669–105679. <https://doi.org/10.1109/ACCESS.2019.2932257>
- Xin, J., Cheng, H., & Ran, B. (2021). Visual servoing of robot manipulator with weak field-of-view constraints. *International Journal of Advanced Robotic Systems*, 18(1), 1–11. <https://doi.org/10.1177/1729881421990320>
- Xu, S., Chou, W., & Dong, H. (2019). A robust indoor localization system integrating visual localization aided by CNN-based image retrieval with Monte Carlo localization. *Sensors*, 19(2), 1–19. <https://doi.org/10.3390/s19020249>
- Xu, Y., Shmaliy, Y. S., Li, Y., Chen, X., & Guo, H. (2019). Indoor INS/LiDAR-based robot localization with improved robustness using cascaded FIR filter. *IEEE Access*, 7, 34189–34197. <https://doi.org/10.1109/ACCESS.2019.2903435>
- Xu, S., Wang, J., Shou, W., Ngo, T., Sadick, A. M., & Wang, X. (2020). Computer vision techniques in construction: A critical review. *Archives of Computational Methods in Engineering*, 28(5), 3383–3397. <https://doi.org/10.1007/s11831-020-09504-3>
- Xu, Z., Zhou, X., & Li, S. (2019). Deep recurrent neural networks based obstacle avoidance control for redundant manipulators. *Frontier in Neurorobotics*, 13(47), 1–13. <https://doi.org/10.3389/fnbot.2019.00047>
- Yang, L., Qi, J., Song, D., Xiao, J., Han, J., & Xia, Y. (2016). Survey of robot 3D path planning algorithms. *Journal of Control Science and Engineering*, 2016, 1–23. <https://doi.org/10.1155/2016/7426913>
- Zafar, M. N., & Mohanta, J. C. (2018). Methodology for path planning and optimization of mobile robots: A review. *Procedia Computer Science*, 133, 141–152. <https://doi.org/10.1016/j.procs.2018.07.018>
- Zake, Z., Chaumette, F., Pedemonte, N., & Caro, S. (2020). Robust 2 1/2D visual servoing of a cable-driven parallel robot thanks to trajectory tracking. *IEEE Robotics and Automation Letters*, 5(2), 660–667. <https://doi.org/10.1109/LRA.2020.2965033>
- Zhang, Z. (2000). A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11), 1330–1334. <https://doi.org/10.1109/34.888718>
- Zhang, K., Chaumette, F., & Chen, J. (2019). Tensor-based 6-DOF visual servoing. *The International Journal of Robotics Research*, 38(10–11), 1208–1228. <https://doi.org/10.1177/0278364919872544>
- Zhang, Y., Li, S., Liao, B., Jin, L., & Zheng, L. (2017). A recurrent neural network approach for visual servoing of manipulators. In 2017 IEEE international conference on information and automation (ICIA), Jul 18–20, (pp. 614–619). <https://doi.org/10.1109/ICInFa.2017.8078981>
- Zhang, H. Y., Lin, W. M., & Chen, A. X. (2018). Path planning for the mobile robot: A review. *Symmetry*, 10(10), 1–17. <https://doi.org/10.3390/sym10100450>
- Zhang, H., Tan, J., Zhao, C., Liang, Z., Liu, L., Zhong, H., & Fan, S. (2020). A fast detection and grasping method for mobile manipulator based on improved faster R-

CNN. *Industrial Robot: The International Journal of Robotics Research and Application*, 47(2), 167–175.
<https://doi.org/10.1108/IR-07-2019-0150>
Zheng, W., Wang, H. B., Zhang, Z. M., Li, N., & Yin, P. H. (2019). Multi-layer feed-forward neural network deep

learning control with hybrid position and virtual-force algorithm for mobile robot obstacle avoidance. *International Journal of Control, Automation and Systems*, 17(4), 1007–1018. <https://doi.org/10.1007/s12555-018-0140-8>



© 2022 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.

You are free to:

Share — copy and redistribute the material in any medium or format.

Adapt — remix, transform, and build upon the material for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made.

You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

No additional restrictions

You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.



Cogent Engineering (ISSN: 2331-1916) is published by Cogent OA, part of Taylor & Francis Group.

Publishing with Cogent OA ensures:

- Immediate, universal access to your article on publication
- High visibility and discoverability via the Cogent OA website as well as Taylor & Francis Online
- Download and citation statistics for your article
- Rapid online publication
- Input from, and dialog with, expert editors and editorial boards
- Retention of full copyright of your article
- Guaranteed legacy preservation of your article
- Discounts and waivers for authors in developing regions

Submit your manuscript to a Cogent OA journal at www.CogentOA.com

