JOURNAL OF MECHANICAL ENGINEERING AND SCIENCES (JMES)
ISSN: 2289-4659 e-ISSN: 2231-8380
VOL. 16, ISSUE 3, 9142 – 9156
DOI: https://doi.org/10.15282/jmes.16.3.2022.15.0724

REVIEW ARTICLE

Review of key technologies for warehouse 3D reconstruction

Y. N. Hao[1,2,*], Y. C. Tan[1], V. C. Tai[1], X. D. Zhang[1,3], E. P. Wei[1,4] and S. C. Ng[5]

[1] Centre for Modelling and Simulation, Faculty of Engineering, Built Environment and Information Technology, SEGI University, 47810 Petaling Jaya, Selangor, Malaysia.
Phone: +60361451777; Fax.:+60361451666
[2] Department of Electrical Engineering, Taiyuan Institute of Technology, No.31 Xinlan Road, Taiyuan, Shanxi 030008, China.
[3] Department of Automation Engineering, Henan Polytechnic Institute, No. 1666, Du Shi Road, Nanyang Henan, 473000, China.
[4] Department of Navigation, Shandong Jiaotong University, No.1508 Hexing Road, Weihai, Shandong,264200, China.
[5] Faculty of Arts and Science, International University of Malaya-Wales, 50480 Kuala Lumpur, Malaysia.

ABSTRACT – Most of the current warehouse management system is made out of two-dimensional (2D) plane schematic warehouse, which brings a lot of inconvenience to warehouse management, including the warehouse data, storage of goods, location search, inventory, etc. 3D warehouse model began to attract attention as it can provide more intuitive view of warehouse-related information. This paper aims to review and investigate the current key technologies used for 3D modeling of warehouse system. This paper reviewed the method of 3D view reconstruction of the warehouse management system, including the active and passive, active and passive fusion methods, and makes a detailed comparison between the active and passive methods. It was found that different methods were applied to reconstruct the 3D view of warehouse, each with its own advantages and disadvantages.

ARTICLE HISTORY
Received: 03rd March 2022
Revised: 17th March 2022
Accepted: 11th Aug. 2022

KEYWORDS
3D reconstruction
Warehouse management system
Binocular
Stereo vision
Fusion

INTRODUCTION

Traditional warehouse management system (WMS) is a management software that provides functionality to handle complex warehouse processes such as transaction of goods, record of inventories, monitor and control of workflow etc., through the use of computer technology [1]. Warehouse storage information relies on manual entry and modification and cannot be updated in a timely manner, nor can it accurately determine the inventory locations, making it inefficient. Figure 1 shows the typical warehousing activities in a warehouse equipped with traditional WMS.

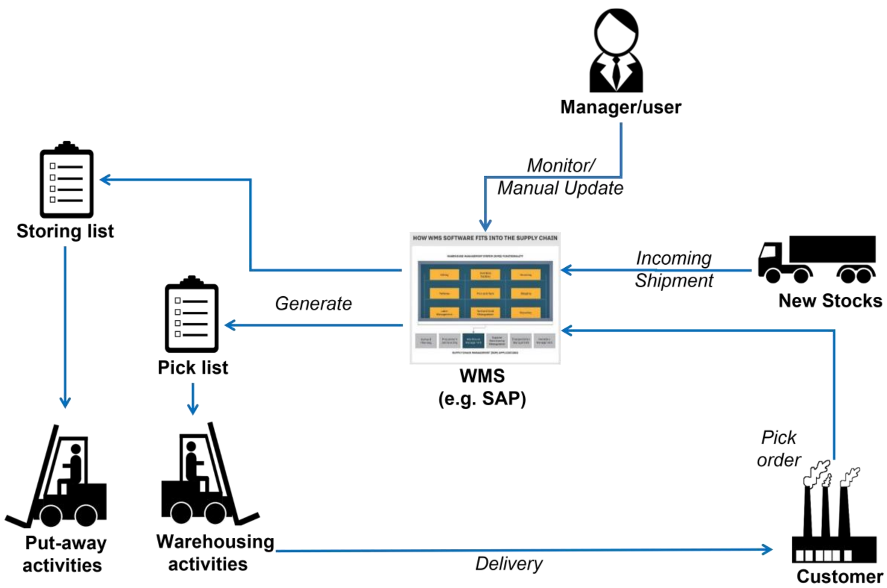

Figure 1. Traditional warehouse management system

On par with the paradigm shift in the manufacturing domain, Logistics 4.0 has been introduced to keep up with the industrial revolution [2-3]. However, Small and Medium-sized Enterprises (SMEs) find it difficult to adopt the Logistics 4.0 concept due to high investment costs [4-5]. Most SMEs are still relying on labors to perform logistics activities in the production warehouse. Non-transparent warehouse operation and inevitable human error in the logistics activities will jeopardize the entire manufacturing processes downstream. For example, in a large automobile manufacturing plant,

*CORRESPONDING AUTHOR | Y. N. Hao | ✉ yananhao2020@163.com

misplaced items and containers may cause significant amount of resource wastage [6]. Automobile manufacturers spend millions of dollars each year just to retrieve and to replace missing containers in a plant. Operational uncertainties from both demand and supply sides are also associated with deleterious impact on logistics services that value high efficiency and low cost. As SMEs are the backbone of economy [7-8], there is an imminent need to provide them a more economic and effective means to solve the logistics problems.

Difficulties of managing inventory in large complex operational environments originate from randomness and various intertwining business processes [9]. It is hard to monitor, collect, and process necessary information and to make real-time decisions based on seemingly random information [10]. Traditional warehouse management systems require operators to update the logistics processes periodically using KANBAN systems [11]. Not only this method is laborious and prone to errors, the management personnel have to be present physically in the warehouse to oversee the logistics operations. Inventory accuracy, space utilization, process management and picking optimization are the major challenges in modern warehouse management [12-13].

Most of the current WMS is made of two-dimensional (2D) plane schematic warehouse, which brings a lot of inconvenience to warehouse management, including the warehouse data, storage of goods, location search, inventory, etc. Based on these problems, 3D warehouse model began to attract attention. 3D warehouse can provide intuitively view warehouse-related information. However, most of the current available 3D models require manual entry of warehouse information to develop the warehouse 3D view. Due to large number of manual data collection, it increases manpower and human error. Although data collection of complex warehouses using 3D scanning, Lidar and other technologies are used to collect 3D information of the warehouse and the construction of 3D model is swift, but these instruments are expensive and cannot be commonly used in SMEs. Computer vision was introduced into warehouse management, using depth cameras, such as ZED binocular camera and special algorithms to construct 3D view of the warehouse and goods [14-17].

WMS is a real-time computer software system, it can realize the storage position, goods according to the specific storage business rules and logic efficient and accurate management of orders and personnel to realize the cost reduction and efficiency increase of storage operations [18].

The world is embarking on revolutionizing the manufacturing industry once more. Known as Industry 4.0, its ultimate goal is to create smart-factories through the use of cyber-physical systems, internet-of-things (IoT), cloud computing, and artificial intelligence (AI), enabling manufacturers to become flexible and better adapt to market changes. Geek+ and Hik Vision introduced AGV intelligent robots to replace manual picking, handling, and sorting, which greatly improved the efficiency of warehousing and reduced labour costs [19]. With the aim of improving the efficiency of warehousing operations, intelligent or smart WMS that incorporates IoT technology with advanced data analytics have been proposed to handle the rising complexity and demands of logistics processes [20-22]. Chen et al. integrated the lean production concept and RFID technology to a distribution warehouse management, significantly shortening the total operation time of the warehousing activities [23]. Laxmi & Mishra pointed out that IEEE 802.11 WLAN network has higher transmission power and higher data rate transfer, making it a good solution to use with RFID to construct an intelligent WMS [24]. Leng et al. proposed a digital twin system (DTS) to map the physical warehouse to the cyber model, in which a joint optimisation is implemented to optimise stacked packing and storage assignment activities using the real-time data [25]. Li et al. proposed the use of 2D and 3D image data from Kinect sensor to monitor warehouse order picking operations [26]. Zou et al. presented an indoor localisation and 3D scene reconstruction modelling method using the Microsoft Kinect sensor[27]. Li proposed drone vision technology to accomplish 3D stereo reconstruction and object recognition [28]. Rashid et al. proposed an RFID-based design solution for automated warehouse management, as well as the use of cloud storage technology for target detection and remote access to data [29]. Lee et al. proposed an IoT based WMS in the face of increasingly complex and diverse customer orders, where warehouse productivity, picking accuracy and efficiency can be improved and robust to order variability [30]. Based on RFID technology, Harry et al. designed and implemented a Resource Management System (RFID-RMS) designed to help users retrieve and analyse useful knowledge from a case-based data warehouse to select the most appropriate resource use package for processing warehouse operations orders [31].

This paper aims to review and investigate the current key technologies used for 3D modeling of warehouse system.

3D RECONSTRUCTION

3D reconstruction is the process of generating the geometry and appearance of real objects using computer vision and computer graphics. Building realistic 3D models of real environments is an important task in computer vision [32].

Accurate indoor scene modelling, as well as fast, low-cost, accurate real-time collection of warehouse visualization data, help to promote the progress tracking and management efficiency of warehouse management. Isgro et al categorized 3D reconstruction into active and passive types. Active type requires a structural light source to be transmitted into the scene and detect the target location by calculating and extracting the projection information of the light source from the scene. Passive type generally uses the surrounding environment such as the reflection of natural light, using the camera to obtain images, and through a specific algorithm to calculate the 3D spatial information of the object [33].

At present, 3D warehouse reconstruction is often implemented by visual and laser. Generally, depth cameras are used to directly obtain depth information, or binocular cameras are used to obtain depth information through calculation, and laser scanning is used for 3D warehouse reconstruction. Although the detection scope is limited, its working principle is more intuitive [34-35].

Active Visual 3D Reconstruction

Active vision-based 3D reconstruction technology has a variety of methods, such as 3D laser scanning / Lidar, TOF (Time of Flight) and structural light vision. This technology mainly uses the optical principle to obtain a data point cloud for warehouse 3D reconstruction.

3D Laser Scanning / Lidar

3D laser measurement technology can quickly and accurately determine the spatial location information of real objects and provide advanced means for a wide range of 3D reconstruction of large or complex geometrical entities [36]. 3D laser scanning uses the Lidar to project laser beams onto the surface and uses triangulation to measure the reflected laser beams for reconstructing of a dense 3D point cloud. Although commercial 3D laser scanners have very high accuracy, the high price greatly limits their application [37].

The laser scanning data process is shown in Figure 2. The point cloud data is first obtained by the laser scanning method, then registration with the original obtained data is conducted to obtain the registered point cloud data, and finally the point cloud data acquired is carried out for a series of processing, so as to obtain the 3D model of the target object [38].

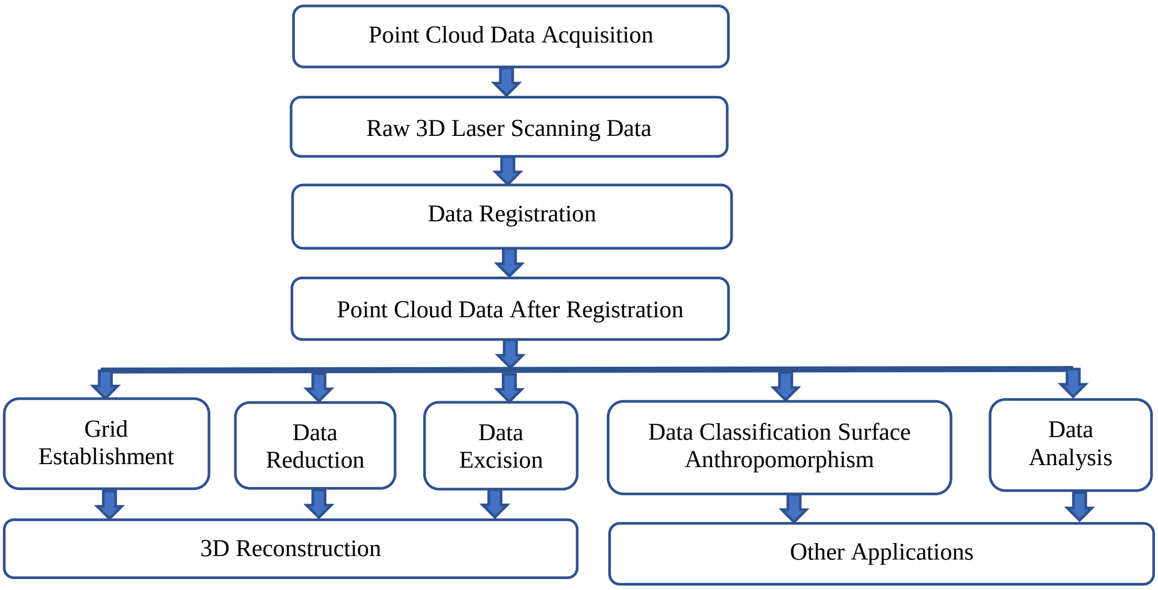

Figure 2. The process of laser scanning data processing [38]

3D laser scanning technology is to represent the information on the surface of the object in the form of point cloud data by laser, and then go through the process of segmentation, denoising and streamlining of the point cloud data, and finally reconstruct the 3D image of the object for modelling. It is a continuous process of data acquisition and data processing, through the point cloud map composed of 3D spatial coordinate values with a certain resolution to express the results of the system's sampling of the surface of the object to be measured [39].

Lidar (Light Detection and Ranging) is the abbreviation for laser scanning and detection system, which mainly consists of a laser transmitter, receiver, processing system and display system. The raw data acquired directly by Lidar through scanning is called depth data. When scanning a target with Lidar, the measurement period of a single point is very short because the laser propagates at the speed of light, and the laser transmitter and receiver can be controlled by a control device to rotate at a certain frequency, thus measuring the Lidar distance to the target within a certain angular range. Scanning in this way can obtain a large range of target 3D data, and the large collection of target data points represented in 3D vector form is called a point cloud [40].

TOF (Time of Flight)

TOF is a kind of active ranging technology. The basic principle is to use the instrument to emit a laser or ultrasound every other time on the measured object, and then to accept the reflected light after meeting the object, and to calculate the depth information of the photographed object by calculating the time difference of the light and the reflection [41]. The principle is shown in Eq. (1) [42].

$$d = \frac{n + \frac{\varphi}{2\pi}}{2}\lambda \tag{1}$$

Among, λ represents the wavelength of the pulse, n denotes the number of wavelengths, φ represents the phase when the pulse returns, d denotes the distance between the object from the emission. TOF camera differs from conventional 3D laser technology in that it can send multiple light pulses at the same time, resulting in an image of all objects within its range. However, the principle of acquiring spatial target points is different from that of binocular system, as binocular measurement is based on the triangulation principle after left-right matching, whereas TOF acquisition is based on the round-trip time of the incident and reflected light [43].

The TOF camera is a new hardware measurement device that can obtain distance data information and gray scale value in real time Moreover, the TOF camera is a face-array-type of active imaging camera that measures the information of the whole scene without scanning, and does not rely on ambient light. The TOF camera is used in the field of 3D reconstruction to calculate the 3D data information of the target object, and then the 3D data information obtained by the TOF camera is used to complete the 3D reconstruction of the warehouse through certain 3D reconstruction methods [44-45].

Structure Light Vision

In the binocular stereo vision measurement, when the optical projector is replaced by one camera, the optical projector projects a certain light mode, such as the light plane, cross light plane, network beam, etc., to restrict the position of the scene object in space, unique coordinate value of the point on the scene object can be obtained, thus forming the structural light 3D visual measurement [46] as shown in Figure 3.

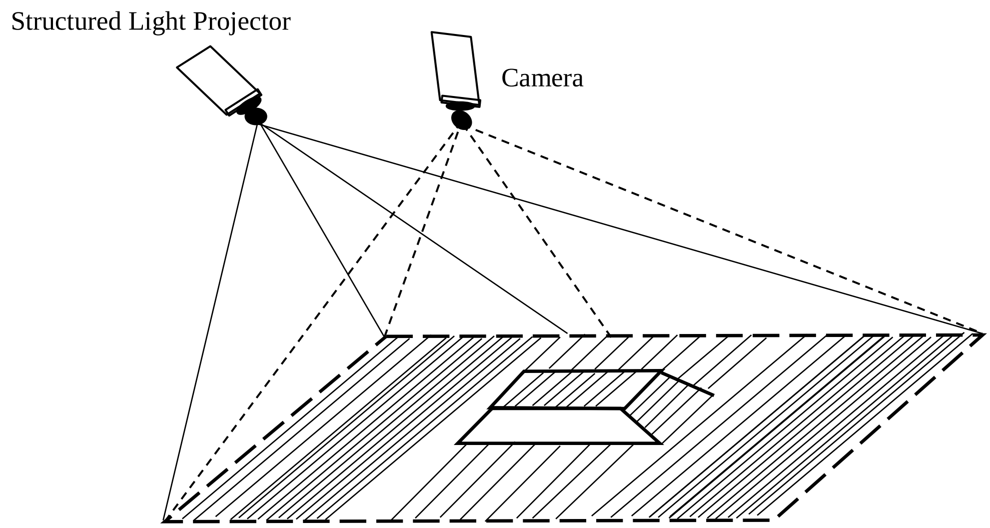

Figure 3. Principles of three-dimensional visual measurement of structured light [46]

Structured light 3D reconstruction methods are divided into line and surface structured light. The line structured light method uses a structured light projector (e.g. laser projector) to project a single line of stripes onto the scene to be reconstructed. The camera acquires the image of the scene containing the single line of stripes and calculates the 3D coordinates of the scene point at which the stripes are projected by means of the position relationship between the line structured light and the camera. Stripe extraction is a fundamental step in the line structured light 3D reconstruction method. Edge extraction, contour extraction and model fitting methods are used in turn to improve the accuracy and speed of the structured light stripe extraction results. The line structured light streak image is fast and computationally efficient, but each shot can only reconstruct the field point corresponding to the line structured light streak, which usually requires multiple scans from different views to complete the scan of the whole target.

Surface structured light 3D reconstruction uses a structured light projector (e.g. a projector) to project a coded (temporally and spatially coded) pattern or stripe onto the surface of the target to be reconstructed. The camera acquires the scene image containing the projected pattern, decodes the pattern information through image processing to obtain the depth information of the target surface, and then calculates the 3D information of the target to be measured [46-47].

Passive Vision 3D Reconstruction

Passive vision-based 3D reconstruction technology is much simpler and more convenient compared to active vision-based 3D reconstruction technology. This method does not need to directly control the use of light sources. Instead, the information in the photographs is analyzed and a 3D model of the actual object is recovered by means of specific algorithms. This technique is relatively simple to operate, and does not require high lighting conditions. The cost required is relatively small, and the feasibility is relatively high. Complex and difficult scenes can be reconstructed in 3D.

Depending on the number of cameras, the 3D reconstruction techniques of passive vision can be divided into monocular, binocular, and multi vision [48].

Monocular Vision

The process of camera photography is the projection of a 3D object in space onto a 2D image. The imaging principle of the camera is usually described by mathematical models, such as the small-aperture model and the binocular model. In the imaging principle model, there are four reference coordinate systems, which are the important groundwork for the subsequent calculation of the 3D coordinates of the object from the 2D coordinates on the image. Pinhole imaging model is the simplest and currently the most widely used model. The specific process of camera imaging can be explained through the analysis of pinhole imaging model as shown in Figure 4.

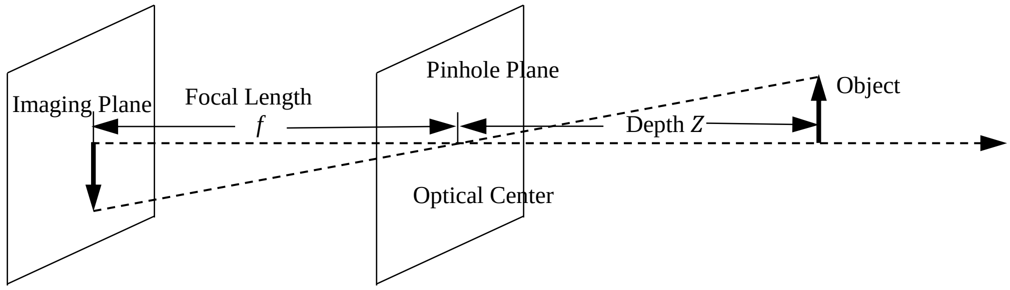

Figure 4. Pinhole imaging model [49]

The pinhole imaging model mainly consists of four parts: pinhole plane, imaging plane, optical center and optical axis. Where, f is the focal length of the camera and Z is the linear distance from the camera's optical center to the 3D object. In the imaging model of the camera, four coordinate systems need to be established to represent the transformation relationship in the imaging process, which are: pixel coordinate system (u,v), image coordinate system (x,y), camera coordinate system (X_c,Y_c,Z_c) and the world coordinate system(X_w,Y_w,Z_w). From the pinhole imaging model, it can be seen that the camera imaging is the transformation of the world coordinate system to the image coordinate system. The conversion relationship between them is described in Eq. (2).

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{dx} & 0 & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha_x & 0 & u_0 & 0 \\ 0 & \alpha_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \tag{2}$$

Where $\alpha_x = f/d_x$ *and* $\alpha_y = f/d_y$ are the normalized focal length, the intrinsic parameter of the camera include $\alpha_x, \alpha_y, u_0, v_0$ of which is only related to the device's own structure. The extrinsic parameter matrix is determined by the rotation matrix R and the translation vector t [50-51].

The monocular vision method is a method of 3D reconstruction with a single camera. The images captured can be single or multiple images from a single viewpoint or multiple images from multiple viewpoints. The former uses 2D features (represented by *X*) to calculate the depth information of a spatially 3D object using a specific algorithm. These 2D features include lightness, darkness, texture, focus, contours, etc., and are therefore also known as shape from *X* methods. This type of method has a simple device structure and can reconstruct a 3D model of an object using a single image or a few images. The disadvantage is that the required conditions are idealized and the reconstruction is generally effective. The latter method matches the same feature points in multiple images. These matching conditions are used to obtain spatial 3D point coordinate information through specific algorithms, thus realizing 3D reconstruction of the object. This method can meet the needs of large-scale scene 3D reconstruction, and the 3D reconstruction effect of multiple images is better. The disadvantages are such as the amount of operations and computation are relatively large and the reconstruction time is longer [52]. 3D reconstruction flow based on monocular vision is shown in Figure 5.

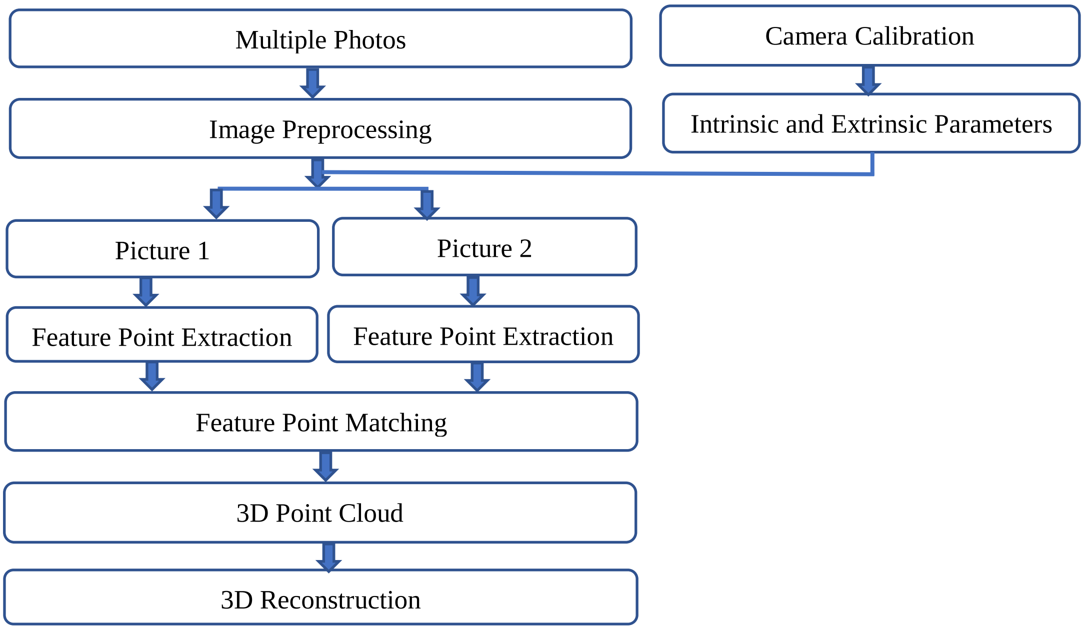

Figure 5. The monocular vision 3D reconstruction flow [52]

Binocular Vision

With the development of laser scanner, the warehouse three dimensional shape information can be accurately captured. However, the optical parts are usually very expensive which makes it difficult to be widely applied. So the 3D reconstruction method based on binocular vision become more popular. A binocular vision system uses two cameras to capture two images of the same scene under realistic lighting conditions. Since the two images are taken from two different angles, 3D information can be recovered from the 2D image of the target object. Schematic diagram of stereo vision is shown in Figure 6 [53]. O_1,O_2 are two cameras, P is a certain point of space object, P point in the left and right imaging plane is P_1 and P_2, respectively. The corresponding points $P_1$and P_2 are obtained through a specific algorithm, which is the focus of stereo matching, and then find the three-dimensional coordinates of space point P according to the triangle PP_1P_2.

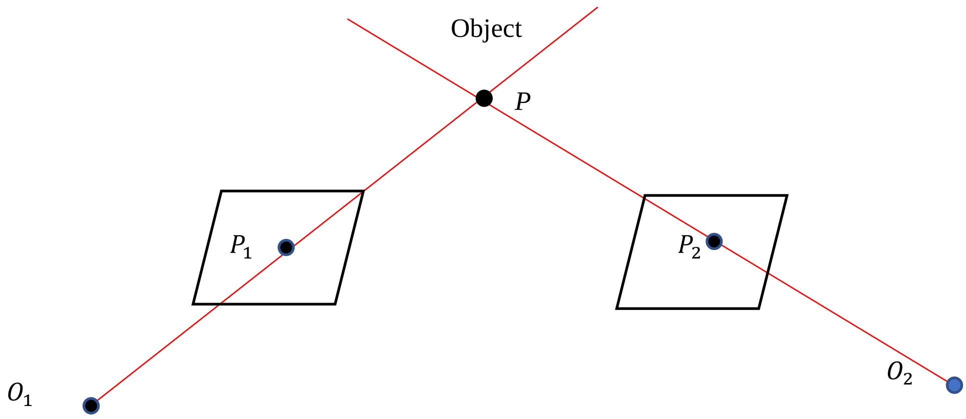

Figure 6. Schematic diagram of stereo vision [53]

The mathematical model of the parallax of the left and right image is shown in Figure 7 [54]. O_l and O_r are the two camera centers, where f is the focus length and b is the baseline and the two camera center point O_lO_r connection. P is a point of space using (X Y Z) coordinates system. The 3D coordinates of the P points can be found based on the properties of the similar triangles, as shown in Eq. (3). Where d is the parallax of the left and right image, $d = x_l - x_r$ [55].

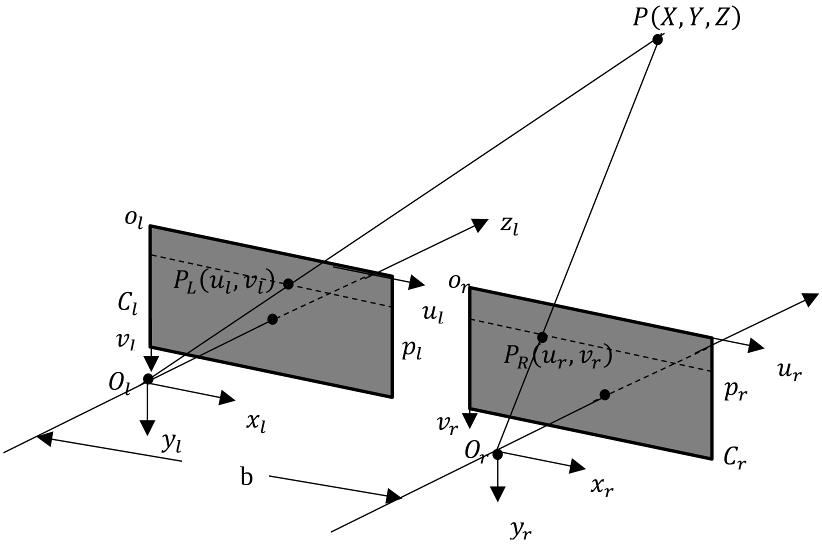

Figure 7. Mathematical model of disparity image in left and right [54]

$$\begin{cases} X = {u_l . B}/{d} \\ Y = {v_l . B}/{d} \\ Z = {f . B}/{d} \end{cases} \quad \frac{u_r - u_l + b}{b} = \frac{Z - f}{z} = \frac{X - x_l}{X}, \frac{f}{Z} = \frac{v_l}{Y} \tag{3}$$

Binocular stereo vision uses two cameras with the same parameters to capture the same object from two viewpoints to obtain two views. It triangulates and calculate the position deviation between the images (parallax) to obtain 3D information about the object. A process similar to the perception process of human vision. A complete stereo vision system can be divided into five parts: image acquisition, camera calibration, feature extraction, stereo matching and 3D reconstruction [56].

Multi-Vision

Multi-eye stereo vision uses of more than two cameras to acquire multiple images of the same target from different viewpoints of the same target. It is an extension of binocular stereo vision. It has the same basic principle as binocular stereo vision. The images of the object's surface are synchronously acquired from different viewpoints by multiple cameras and these images can be processed to reconstruct the model. The main difference is that multi-vision method increases cameras to obtain more viewpoints, which can resolve the problem of limited space and occlusion when large-size and complex objects are reconstructed. Multi-vision system model is as shown in Figure 8 [57].

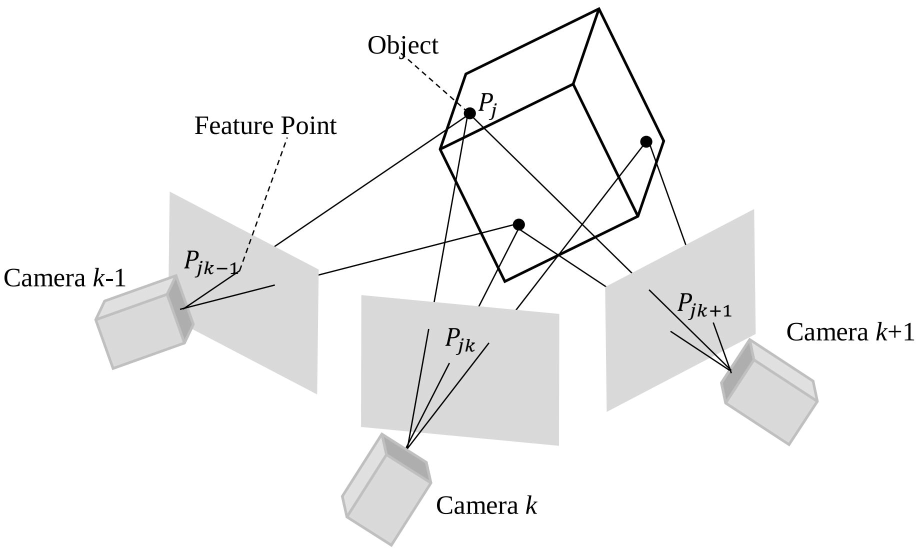

Figure 8. Diagram of muti-view reconstruction [58]

The main problems of binocular vision method in the 3D reconstruction process are: firstly, the presence of duplicate or similar features in the image is prone to the generation of false targets; secondly, when using epipolar constraints, the edges parallel to the epipolar lines are prone to blurring; thirdly, if the baseline distance increases, making the occlusion serious, the number of spatial points that can be reconstructed decreases, and at the same time, due to the increase in the parallax range, resulting in a larger search space in which the possibility of incorrect matching in a larger search space also increases. In response to these problems, a triocular vision method is proposed, and its basic idea is to provide additional constraints by adding one camera as a way to avoid the binocular vision problems [59].

Fusion Vision Methods

The combination of active and passive visual measurement is by adding specific light source mode in the passive visual measurement system. With the help of the projected light source mode, the matching problem in passive binocular vision can be solved, so as to improve the matching accuracy and measurement accuracy.

The active camera has a low resolution, for example, the TOF has a resolution of only 320×240 or less, it is unable rebuild fine-object structures. However, the passive 3D reconstruction method can restore the more refined structure of the high-texture area, and it can easily obtain high-resolution maps in high-speed detection. However, the binocular stereo matching algorithm has high requirements for the scene, and the weak texture area will produce many mismatches. Thus, TOF camera can combine with a passive camera to improve the resolution, moreover, active sensors are able to sense depth even in non-textured regions, which compensates for the disadvantage of low robustness of active cameras in weak texture regions [60].

Comparison of the active and passive sensors are listed in Table 1. As can be seen, active and passive sensors complement each other [61].

Table 1. Comparison of active and passive sensors [61]

Sensor	Resolution	Highly-textured Region	Non-textured
Active	Low	Non-robust	Robust
Passive	High	Robust	Non-robust

Lidar + Stereo Vision

Both Lidar and binocular stereo vision are capable of recovering 3D warehouse. Binocular stereo vision uses the relationship between feature points corresponding to the left and right views and uses specific algorithms to recover 3D warehouse. However, as it uses images as input data, it is affected by changes in ambient lighting and weakly textured areas that are difficult to match. Lidar measures the 3D information of a scene directly by emitting a laser. This approach is accurate and less affected by changes in the extrinsic environment but has the disadvantage of sparse data. The fusion of Lidar and stereo vision can therefore compensate for their respective disadvantages [62].

The fusion of Lidar and binocular stereo vision can be divided into two main types: post-fusion and pre-fusion. Post-fusion refers to the fusion of 3D data measured by Lidar and 3D data measured by binocular stereo vision at the data level to enhance the final system output of 3D data. This fusion is generally direct to different areas of the scene according to the strengths and weaknesses of the two data to directly fuse, the amount of data and complex calculations. On the contrary, pre-fusion is the fusion of Lidar data as a priori information into the stereo vision calculation method to improve the parallax map obtained by stereo vision. This fusion uses the Lidar a priori information to constrain the stereo matching algorithm, the data volume is small and can improve the speed of stereo matching [63-65]. Figure 9 shows data fusion flowchart.

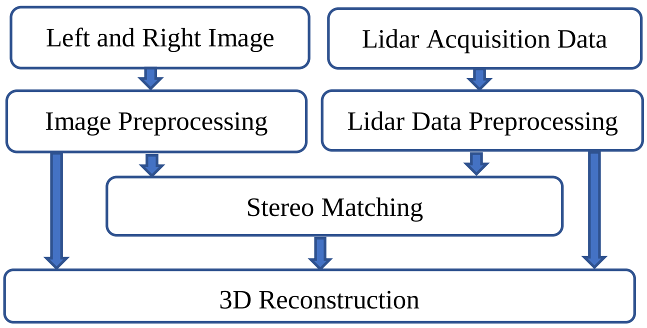

Figure 9. Data fusion flowchart [62]

Combining Lidar and dual camera for 3D reconstruction has four main modules: image preprocessing module, radar data preprocessing module, three-dimensional matching module and 3D reconstruction module.

TOF + Stereo Vision

Binocular stereo vision is a classical passive technique and is the most used stereo vision technique. Binocular stereo matching is capable of obtaining dense, high resolution, high accuracy parallax maps but it has its own limitations such as very complex algorithms, susceptibility to interference from environmental factors, dependence on ambient light sources, neither matching algorithm is a good solution to the problems of occlusion, lack of texture features. TOF technology can directly acquire depth information, can obtain a denser 3D point cloud than structured light, and has better real-time performance, lower cost, lower body performance and power requirements, and no dependence on the surface texture of the object. However, its disadvantages are such as unreliable results at the edge of the field of view, relatively low resolution, (even 320 × 240 or less) and lack of depth measurement at closer distances. Therefore, based on the complementary properties between the TOF depth camera and binocular stereo matching, the combination of the TOF depth camera and binocular stereo technology can result in a more robust parallax map, ultimately extending the applicability of the camera's 3D reconstruction to obtain a more complete and highly accurate 3D point cloud, and thus providing a better warehouse 3D reconstruction [66].

The combination of TOF and binocular stereo vision is of great practical importance, as both methods have their advantages and disadvantages, and it is difficult to use a single technique for all scenes. By combining the two, the advantages of each are fully exploited and better reconstructions can be obtained. The combination of the two methods improves the warehouse reconstruction efficiency, reconstruction accuracy and algorithm robustness of the 3D reconstruction method while ensuring the reconstruction accuracy. Figure 10 shows the flow of the fusion method [67].

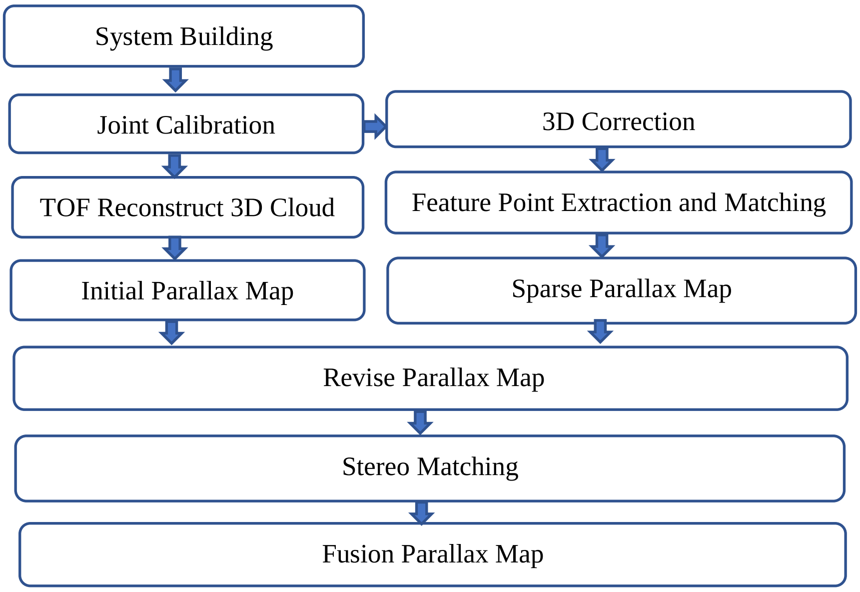

Figure 10. Flowchart of fusion method [67]

Structure Light + Stereo Vision

Binocular is a stereo matching of the picture taken and 3D reconstruction, but when texture, geometry and other information are missing in the scene, Binocular stereo vision cannot effectively perform 3D reconstruction. Structural light technology through the optical projector projection pattern, increase the features within the scene, can reconstruct the three-dimensional scenes with missing features [68].

The structured light is applied to the surface of the object to be measured, and the distorted image of the object surface is captured by two cameras, which are then processed to obtain the information of the corresponding spatial points of the light bars, and finally combined with the intrinsic and extrinsic parameters of the two cameras to recover the 3D spatial information of the object surface projected by the structured light. The structured light binocular vision system consists of two optical cameras, a structured light source, a computer, and other hardware [69]. Structural light binocular visual measurement model is shown in Figure 11.

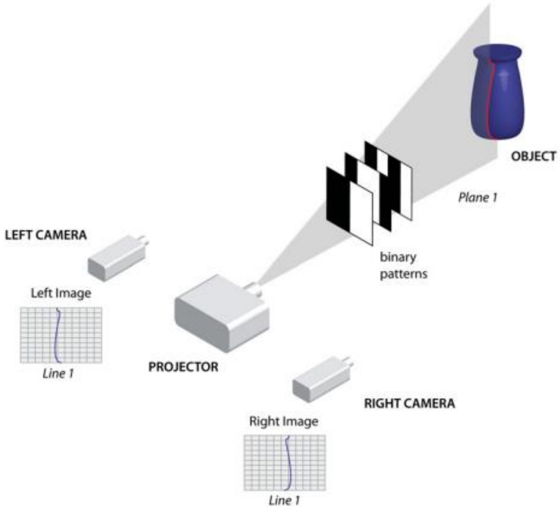

Figure 11. Structural light binocular visual measurement model [70]

DISCUSSION

Traditional warehouse reconstruction lacks three-dimensional and intuitive. 3D warehouse reconstruction provides a more intuitive view of the warehouse, including shelves and goods identification. Only through this method, can we facilitate the 3D visual management of the warehouse, and save manpower and material resource. The current mainstream 3D reconstruction methods are reviewed, including active, passive, and fusion methods of 3D reconstruction methods. Active method including 3D scanning, TOF camera and structure light. Passive methods include monocular, binocular and multi-ocular (with only trinocular 3D reconstruction). Table 2 shows the comparison of the 3D reconstruction methods.

This paper reviewed both active and passive fusion 3D reconstruction methods such as laser scanning + stereo fusion, TOF + stereo fusion. Laser scanning method is mainly used to scan the 3D point cloud data on the object surface for warehouse reconstruction, the method not only able to establish a simple shape object 3D model, but can also generate irregular object 3D model and generate 3D model accuracy. However, laser scanning method cannot identify the shelves and goods in the warehouse, which needs to combine stereo image acquisition, in order to achieve the purpose of identification and reconstruction [71].

The application of TOF camera in the field of 3D reconstruction is mainly by calculating the 3D data information of the camera system imaging model and distance data information, and then using the 3D data information to complete the 3D reconstruction of the warehouse. TOF camera is easy to be affected by noise interference and low depth map resolution. However, stereo vision can obtain high resolution depth map. Based on the complementary advantages of TOF and stereo vision, obtaining a better effect 3D view. For example, by using TOF method in weak texture or repeated texture area, and stereo vision in complex texture area to obtain high quality depth map. Combine the two methods to complete the shelves and cargo identification and reconstruction, get a better effect warehouse 3D view [72].

Binocular stereo equipment is simple and the effect is good, but the matching algorithm is complex. Structure light method is not affected by light change and texture, and has high robustness. Combining the advantages of the two methods, structural light was added to the passive binocular measurements, the main purpose is to reduce mismatching, to enhance its robustness and accuracy and to improve warehouse reconstruction accuracy [73].

Table 2. Comparison of 3D reconstruction methods [35,44,61,71,74-78]

3D Reconstruction	**Active Method**			**Passive Method**		
Methods	3D Scanning	TOF	Structure Light	Monocular	Binocular	Multi-Vision (Trinocular)
Operational Principle	Information on the object surface by laser is presented in the form of point cloud data	Reflect light time difference	Encode structural light	Individual cameras are taken from different angles for 3D reconstruction through some column algorithms	Simulated human eyes took multiple sets of photos from different angles, camera calibration, feature point extraction, stereo matching, and finally 3D reconstruction	Multiple cameras, shot from various different angles

Table 2. Comparison of 3D reconstruction methods [35,44,61,71,74-78] (cont.)

3D Reconstruction	**Active Method**			**Passive Method**		
Measuring Range	Radius is up to 25m	not within 100m	0-10m	0-150m	0-10m	0-10m
Anti-Environmental Light	Higher	Lower	Higher	Lower		
Influence Factors	Blocking problem, only the surface facing Lidar can be scanned	Unaffected by light variations and textures, affected by multiple reflections	Unaffected by light variations and textures, affected by reflections	Easy to light and stripe changes, and unavailable at night		
Power Consumption	High	Very high	Middle	Low		
Image Resolution Ratio	None	Middle	Lower	Higher		
Measurement Accuracy	0.8%	0.01mm-1mm	1%	3%	2%	2%
Advantage	High accuracy. Fast speed, rich details, need no lighting. Indoor and outdoor are suitable	High precision, strong real-time performance	High robustness and high measurement accuracy	Low cost and easy to operate	Adaptable, reliable and flexible	Reduces blind spots in measurements and solves mis-matching phenomena
Disadvantage	Large amount of data and expensive equipment. The detection range is limited. The reconstructed model lacks the semantic information	High hardware cost and slow reconstruction speed	The resolution is low, and the measured 3D point cloud error is large, up to 30cm	Large operation volume. Scale missing. low accuracy	Characteristic extraction and matching were difficult. baseline constraints. Small range measurement	Large amount of data to be processed for real-time. The equipment is complex, with a high cost and a long reconstruction time. Baseline constraints.
Reconstruction Effect	The effect is more meticulous	The effect is better	General	The reconstruction effect is poor which is slightly better with rich image resources.	Better, but the reconstruction effect is significantly lower with a larger baseline	The reconstruction is better than the binocular
Application scenarios	Building surface reconstruction, topographic survey, industrial robots and other high-precision modeling	Cultural relics digitization, human three-dimensional face modeling, etc., is not suitable for outdoor	Rich application scenarios are suitable, indoor and outdoor	Human 3D modeling, large-scale 3D scene reconstruction	Can meet most scenes, more suitable for close range	Autonomous vehicle driving, robot vision, multi-degree-of-freedom mechanical device control

CONCLUSION

With the rapid development of sensor technology and computer vision, the traditional two-dimensional warehouse view can no longer meet people's needs. The three-dimensional reconstruction of the warehouse can be more realistic, more intuitively to view the warehouse shelves and goods location, which is conducive to the improvement of the level of warehouse automation management. In order to realize the more intuitive and efficient warehouse management, it is necessary to introduce 3D reconstruction technology. Current 3D reconstruction methods can generally be divided into active and passive types. Active refers to using light or energy sources such as lasers, acoustic waves, electromagnetic waves, etc. to emit to the target object and receiving the returned light waves to obtain the depth information of the object. The active type consists of 3D laser scanning/Lidar, TOF and structured light. Passive methods can be divided into monocular reconstruction, binocular reconstruction and multi-vision reconstruction according to the number of cameras.

1. 3D laser scanning technology/Lidar can continuously, automatically and quickly scan the surface of the object to be measured to obtain a large amount of 3D coordinate information, i.e. point clouds. The 3D reconstruction of warehouses through this method is effective, with high accuracy and fast reconstruction, More importantly, it can solve the problem of visual 3D reconstruction affected by high light. However it lacks semantic information,a huge amount of data information needs to be processed and the equipment cost is high, therefore it not widely used.
2. TOF is not affected by light and object texture, the software is relatively simple, can be used in both indoor and outdoor. It has high robustness and high measurement accuracy, and is more suitable for long-distance of 3D reconstruction. But the resolution is low, and the measured 3D point cloud error is large. So the reconstruction of WMS reconstruction effect is not meeting expectation.
3. The structural light method is not affected by light change and texture, and the measurement range is within 10 meters, so it is suitable for three-dimensional reconstruction. The reconstruction effect is good, the accuracy is close to 0.01-0.1mm, but the high hardware cost, affected by reflection, slow reconstruction speed, complex algorithm, and high power consumption.
4. Passive 3D reconstruction, whether monocular vision or multi-vision, the image reconstruction process have their own application advantages and disadvantages. Compared to monocular stereo vision, multi-vision stereo vision is more widely used and adopted in practice. Taking into account the cost and complexity of the algorithm, in the 3D reconstruction of the commonly used type of binocular 3D reconstruction, binocular 3D reconstruction cost, low power consumption, accuracy depends on the baseline, close accuracy can reach millimetre, but easily affected by the light, the algorithm is complex, the reconstruction effect is generally, more suitable for close 3D reconstruction.
5. Regardless of the 3D reconstruction method, it has advantages and disadvantages, combining the active and passive 3D reconstruction methods to improve the reconstruction accuracy and reconstruction results. This paper reviewed three fusion modes. Lidar + stereo vision combination, TOF + stereo vision combination, structured-light + stereo vision combination. Because of the low TOF resolution, the combination of the TOF depth camera and binocular stereo technology can result in a more robust parallax map, ultimately extending the applicability of the camera's 3D reconstruction to obtain a more complete and highly accurate 3D point cloud, and thus providing a better 3D warehouse reconstruction. Binocular utilizes the relationship between feature points corresponding to the left and right views, and uses specific algorithms to restore warehouse 3D views, but it is influenced by changes in environmental illumination and weak-texture regions that are difficult to match. Lidar directly measures the 3D information of a scene with an emission laser. The method is less affected by changes in the extrinsic environment, but the disadvantage is that the data is sparse. Therefore, the fusion of Lidar and stereo vision can compensate for their respective shortcomings. What's more, structural light technology artificially increases the features in the scene, which can reconstruct 3D scenes with missing features. So these three fusion methods can reconstruct a better effective 3D view of the warehouse.

ACKNOWLEDGMENTS

The project is funded by the Ministry of Higher Education Malaysia, under the Fundamental Research Grant Scheme (FRGS Grant No. FRGS/1/2018/TK03/SEGI/02/1).

REFERENCES

[1] W. Ding, "Study of smart warehouse management system based on the IOT," *Advances in Intelligent Systems and Computing*, vol. 180, pp. 203-207, 2013, doi: 10.1007/978-3-642-31656-2_30.

[2] J. O. Strandhagen, L. R. Vallandingham, G. Fragapane, and J. W. Strandhagen, "Logistics 4.0 and emerging sustainable business models," *Advances in Manufacturing*, vol. 5, no. 4, pp. 359-369, 2017, doi: 10.1007/s40436-017-0198-1.

[3] L. Barreto, A. Amaral, and T. Pereira, "Industry 4.0 implications in logistics: an overview," *Procedia Manufacturing*, vol. 13, pp. 1245-1252, 2017, doi: 10.1016/j.promfg.2017.09.045.

[4] L. Sommer, "Industrial Revolution-Industry 4.0: Are German manufacturing SMEs the first victims of this revolution," *Journal of Industrial Engineering and Management*, vol. 8, no. 5, pp. 1512-1532, 2015, doi: 10.3926/jiem.1470.

[5] L. D. Gamindo, "The challenges of Logistics 4.0 for the supply chain management and the information technology," Master thesis, Norwegian University of Science and Technology, Trondheim, Norway, 2016.

[6] W. Zhou, S. Piramuthu, F. Chu, C. Chu, "RFID-enabled flexible warehousing," *Decision Support Systems*, vol 98, pp. 99- 112, 2017.

[7] H. Karadag, "The role of SMES and entrepreneurship on economic growth in emerging economies within the post-crisis era: an analysis from Turkey," *Journal of Small Business and Entrepreneurship Development*, vol. 4, no. 1, pp. 22-31, 2016.

A. Issa, D. Lucke, and T. Bauernhansl, "Mobilizing SMEs towards Industrie 4.0-enabled smart products," *Procedia CIRP*, Vol. 63, pp. 670 -674, 2017.

[8] Z. Taha, V. C. Tai, and P. C. See, "Design and manufacture of a miniature UAV using 3D rapid prototyping", *Advanced Material Research*, vol. 308-310, pp. 1425-1435, 2011.

[9] D. Sanders, B. Sanders, D. Ndzi, and N. Bausch, "Using confidence factors to share control between a mobile robot tele-operater and ultrasonic sensors," *IEEE Proceedings of SAI Intelligent Systems Conference (IntelliSys) UK*. ISBN 978-1-5090-6435-9, pp: 1026-1033, 2017.

[10] C. Silva, L.M. Ferreira, M. Thurer, and M. Stevenson, "Improving the logistics of a constant order-cycle kanban system," *Production Planning and Contro*l, vol. 27, pp. 650-659,2016.

[11] T. L. Landers, M. H. Cole, B. Walker, and R.W. Kirk, "The virtual warehousing concept," *Transportation Research Part E*, vol. 36, pp. 115-126, 2000.

[12] S. H. Fung, C. F. Cheung, W. B. Lee, and S. K. Kwok, "A virtual warehouse system for production logistics," *Production Planning and Control*, vol. 16, no. 6, pp. 597- 607, 2005.

[13] W. J. Chen, "Design of a 3-dimensional visual simulation monitoring system for warehouse logistics management," *Network and information*. pp. 421-422+453, 2014,

[14] W. C. Tan, and K. M. Yap, "Exploratory research on application of different vision system on warehouse robot using selective algorithm," In: *Advances in Visual Informatics, IVIC2017. Lecture Notes in Computer Science,* vol. 10645, pp. 290-296, 2017, doi: 10.1007/978-3-319-70010-6_27.

A. Causo, Z. -H. Chong, R. Luxman and I. -M. Chen, "Visual marker-guided mobile robot solution for automated item picking in a warehouse," In: *2017 IEEE International Conference on Advanced Intelligent Mechatronics (AIM)*, pp. 201-206, 2017, doi: 10.1109/AIM.2017.8014018.

[15] Z. -H. Chong, R. Luxman, W. -C. Pang, Z. Yi, R. Meixuan, H. S. Tju, A. Causo, and I. -M. Chen, "An innovative robotics stowing strategy for inventory replenishment in automated storage and retrieval system," In: *2018 15th International Conference on Control, Automation, Robotics and Vision (ICARCV)*, pp. 305-310, 2018, doi: 10.1109/ICARCV.2018.8581338.

[16] S. C. Ng, V. C. Tai, Y. C. Tan, and N. F. A. Rahman, "SFlex-WMS: a novel multi-expert system for flexible logistics and warehouse operation in the context of Industry 4.0," *SHS Web Conf.*, vol. 124, paper no. 10002, 2021, doi: 10.1051/shsconf/202112410002.

[17] Y. Wu, and D. Ge, "Key technologies of warehousing robot for intelligent logistics" *Advances in Social Science, Education and Humanities Research*, vol. 309, pp. 79 – 82, 2019, doi: 10.2991/ismss-19.2019.16.

[18] C. K. M. Lee, Y. Lv, K. K. H. Ng, W. Ho, and K. L. Choy, "Design and application of internet of things-based warehouse management system for smart logistics," *International Journal of Production Research*, vol. 56, no. 8, pp. 2753-2768, 2018, doi: 10.1080/00207543.2017.1394592.

[19] Y. Ding, M. Jin, S. Li, and D. Feng, "Smart logistics based on the internet of things technology: an overview," *International Journal of Logistics Research and Applications*, vol. 24, no. 4, pp. 323-345, 2021, doi: 10.1080/00207543.2017.1394592.

[20] S. Winkelhaus, and E.H. Grosse, "Logistics 4.0: a systematic review towards a new logistics system," *International Journal of Production Research*, vol.58, no. 1, pp. 18-43, 2020, doi; 10.1080/00207543.2017.1394592.

[21] J. C. Chen, C. -H. Cheng, P. B. Huang, K. -J. Wang, C. -J. Huang, and T. -C. Ting, "Warehouse management with lean and RFID application: a case study," *International Journal of Advanced Manufacturing Technology*, vol. 69, pp. 531-542, 2013, doi: 10.1007/s00170-013-5016-8.

A. R. Laxmi, and A. Mishra, "RFID based logistic management system using internet of things (IoT)," *In: 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, pp. 556-559, 2018, doi: 10.1109/ICECA.2018.8474721.

[22] J. Leng, D. Yan, Q. Liu, H. Zhang, G. Zhao, L. Wei, D. Zhang, A. Yu, and X. Chen, "Digital twin-driven joint optimisation of packing and storage assignment in large-scale automated high-rise warehouse product-service system," *International Journal of Computer Integrated Manufacturing*, vol. 34, no. 7-8, pp. 783-800, 2021, doi: 10.1080/0951192X.2019.1667032.

[23] X. Li, I. Y. H. Chen, S. Thomas, and B. A. MacDonald, "Using kinect for monitoring warehouse order picking operations" In: Proceedings of Australasian Conference on Robotics and Automation, 3-5 December 2012, *Victoria University of Wellington*, New Zealand, 2012.

[24] Y. Zou, W. Chen, X. Wu and Z. Liu, "Indoor localization and 3D scene reconstruction for mobile robots using the Microsoft Kinect sensor," *In: IEEE 10th International Conference on Industrial Informatics*, pp. 1182-1187, 2012, doi: 10.1109/INDIN.2012.6301209.

[25] X. H. Zhou, Z. M. Yi, K. H. Huang, and H. Huang, "Survey on path and view planning for UAVs", *Virtual Reality & Intelligent Hardware*, vol. 2, no. 1, pp. 56-59, 2020.

[26] R. Mahmudur. S. M. Abdul Ahad, S. Siddique, and T. Motahar, "Smart warehouse management system with RFID and cloud database," In: 2019 Joint 8th International Conference on Informatics, *Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (icIVPR)*, 2019, pp. 218-222, doi: 10.1109/ICIEV.2019.8858546.

[27] C. K. M. Lee, Y. Lv, K. H. Ng, and W. Ho, "Design and application of internet of things-based warehouse management system for smart logistics," *International Journal of Production Research*, pp. 1-16, 2017, doi:10.1080/00207543.2017.1394592.

[28] H. K. H. Chow, K. L. Choy, W. B. Lee, and K. C. Lau, "Design of a RFID case-based resource management system for warehouse operations," *Expert Systems with Application*, vol. 30, no. 4, pp. 561-576, 2006, doi: 10.1016/j.eswa.2005.07.023.

[29] K. R. Lee, "Accurate, efficient, and robust 3D reconstruction of static and dynamic objects." *Doctor Thesis*, University of California, American, 2014.

[30] F. Isgro, F. Odone, and A. Verri, "An open system for 3D data acquisition from multiple sensor," *Seventh International Workshop on Computer Architecture for Machine Perception (CAMP'05)*, 2005, pp. 52-57, doi: 10.1109/CAMP.2005.13.

[31] L. X. Huan, X. W. Zheng, and J. Y. Gong, "GeoRec: Geometry-enhanced semantic 3D reconstruction of RGB-D indoor scenes," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 186, pp. 301-314, 2022.

[32] C. L. Wen, J. B. Tan, F. S. Li, C. R. Wu, Y. T. Lin, Z. Y. Wang, and C. Wang, "Cooperative indoor 3D mapping and modeling using LiDAR data," *Information Sciences*, vol. 574, pp. 192-209, 2021.

[33] L. Yang, Y. H. Sheng, and B. Wang, "3D reconstruction of building facade with fused data of terrestrial LiDAR data and optical image," *Optik*, vol. 127, no. 4, pp. 2165-2168.

[34] D. Li, L. Xu, X. S. Tang, S. Sun, X. Cai, and P. Zhang, "3D imaging of greenhouse plants with an inexpensive binocular stereo vision system," *Remote Sensing*, vol. 9, no. 5, doi:10.3390/rs9050508, 2017.

[35] Y. S. Yin, and J. Antonio, "Application of 3D laser scanning technology for image data processing in the protection of ancient building sites through deep learning, " *Image and Vision Computing*, vol. 102, 2020, doi: 10.1016/j.imavis.2020.103969.

[36] Y. Pan, Y. Han, L. Wang, J. Chen, H. Hao, G. Q. Wang, Z. C. Zhang, and S. B. Wang. "3D Reconstruction of Ground Crops Based on Airborne LiDAR Technology," *IFAC-PapersOnLine*, vol: 52, pp, 35-40, 2019, doi: 10.1016/j.ifacol.2019.12.376.

[37] S. H. Venugopala. "Comparative study of 3D object detection frameworks based on LiDAR data and sensor fusion techniques," *Journal of Physics: Conference Series,* 2022, doi: 10.1088/1742-6596/2232/1/012015.

[38] S. Vandenberghe, and J. Karp, "Rebinning and reconstruction techniques for 3D TOF-PET," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 569, no. 2, pp. 421-424.

[39] T. X. Zheng, S. Huang, Y. F. Li, and M. C. Feng, "Key techniques for vision based 3D reconstruction-A review," *ACTA AUTOMATICA SINICA*, vol. 46, no. 4, pp. 631-652, doi:10.16383/j.aas.2017.c170502, 2020.

[40] S. Á. Guðmundsson, M. Pardàs, J. R. Casas, J. R. Sveinsson, H. Aanæs, and R. Larsen, "Improved 3D reconstruction in smart-room environments using ToF imaging," *Understanding*, vol. 114, no. 12, pp. 1376-1384, 2010.

[41] G. Alenyà, S. Foix, and C. Torras, "ToF cameras for active vision in robotics," *Sensors and Actuators A: Physical* , vol. 218, pp. 10-22, 2014.

[42] S. Á. Guðmundsson, M. Pardàs, J. R. Casas, J. R. Sveinsson, H. Aanæs, and R. Larsen, "Improved 3D reconstruction in smart-room environments using ToF imaging," *Computer Vision and Image Understanding*, vol. 114, no. 12, pp. 1376-1384,2010, doi: 10.1016/j.cviu.2010.07.011.

[43] J. Geng, "Structured-light 3D surface imaging: a tutorial," pp. 131, 2008.

[44] M. Y. Li, J. Liu, H. M. Yang, W. Q. Song, and Z. H. Yu, "Structured light 3D reconstruction system based on a stereo calibration plate," *Symmetry,* vol. 12, no. 5, pp. 772, 2020, doi: 10.3390/sym12050772.

[45] Z. Z. Kang, J. T. Yang, Z. Yang, and S. Cheng, "A review of techniques for 3D reconstruction of indoor environments," *ISPRS International Journal of Geo-Information,* vol. 9, no. 5, pp.330, 2020, doi: 10.3390/ijgi9050330.

[46] D. D. Lv, and G. E. Jiao "Experiment of stereo matching algorithm based on binocular vision," *Journal of Physics: Conference Series,* vol. 1574, 2020, doi: 10.1088/1742-6596/1574/1/012173.

[47] S. F. Zhuang, X. Zhang, D. W. Tu, C. Zhang, and L. L. Xie, "A standard expression of underwater binocular vision for stereo matching," *Measurement Science and Technology,* vol. 31, no. 11, pp. 5012, 2020, doi: 10.1088/1361-6501/ab94fd

[48] N. Guo, L. Li, F. Yan, and T. Li, "Binocular stereo vision calibration based on constrained sparse beam adjustment algorithm," *Optik*, vol. 208, pp. 1-11. doi: 10.1016/j.ijleo.2019.163917, 2020.

[49] B. Q. Xu , and C. Liu "A 3D reconstruction method for buildings based on monocular vision," *Computer-Aided Civil and Infrastructure Engineering*, 2021, doi: 10.1111/mice.12715

[50] H. Y. Yang, "Binocular stereo vision based three dimensional reconstruction using domain-sized pooling local features," *2nd International Conference on Artificial Intelligence: Techniques and Applications*, pp. 364-368, doi: 978-1-60595-491-2, 2017.

[51] G. Huang, "Binocular vision system for real-time badminton tracking," *Journal of electronic measurement and instrumentation*, vol. 35, pp. 117-123, doi:10.13382 /j.Jemi.B2003616, 2021.

[52] L. E. Ortiz, E. V. Cabrera, and L. M. Goncalves, "Depth data error modeling of the ZED 3D vision sensor from stereolabs," *Electronic Letters on Computer Vision and Image Analysis*, vol.17, no. 1, pp. 1-15, 2018.

[53] Huo G, Wu Z, Li J, and Li S, "Underwater Target Detection and 3D Reconstruction System Based on Binocular Vision," *Sensors*, vol, 18, no. 10, 2018, doi: 10.3390/s18103570.

[54] Z. H. Zhang, W. Liu, G. D. Liu, L. M. Song, Y. F. Qu, X. D. Li, and Z. Z. Wei "Development of three-dimensional vision measurement technology and its application," *Journal of Image and Graphics*, vol. 26, no. 6, pp. 1483-1502, 2021, doi: 10. 11834 /jig. 200841

[55] R. Hartly, and A. Zisserman, "Multiple view geometry in computer vision, "2nd Edition. Cambridge University, England, 2004.

[56] Y. N. Ma, Q. O. Li, J. Xing, G. Y. Huo, and Y. Liu, "An intelligent object detection and measurement system based on trinocular vision," *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 30, no. 3, pp. 711-724, 2020, doi: 10.1109/tcsvt.2019.2897482.

[57] X. Y. Wang, T. P. Lin, X. S. Jiang, K. Xiang, and F. Pan, "Reliable fusion of ToF and stereo data based on joint depth filter," *Journal of Visual Communication and Image Representation*, vol. 74, 2021, doi: 10.1016/j.jvcir.2020.103006.

[58] Q. Yang, K. H. Tan, B. Culbertson, and J. Apostolopoulos, "Fusion of active and passive sensors for fast 3D capture," *IEEE International Workshop on Multimedia Signal Processing*, 2010.

[59] L. Yang, Y. H. Sheng, and B. Wang, 3D reconstruction of building facade with fused data of terrestrial LiDAR data and optical image, *Optik*, vol. 127, no. 4, pp. 2165-2168, 2016, doi: 10.1016/j.ijleo.2015.11.147.

[60] P. Gurram, H. Rhody, J. Kerekes, S. Lach, and E. Saber, "3D scene reconstruction through a fusion of passive video and lidar imagery," *AIPR. 36th IEEE*. pp. 133-138, 2007.

[61] K. D. Kuhnert, and M. Stommel, "Fusion of stereo-camera and pmd-camera data for realtime suited precise 3D environment reconstruction," *Intelligent Robots and Systems. IEEE/RSJ International Conference*, pp. 4780-4785, 2006.

[62] K. M. Nickels, A. Castaño, and C. Cianci, "Fusion of Lidar and stereo range for mobile robots," *Proceedings of ICAR 2003*, pp. 65-70, 2003.

[63] Y. Deng, J. M. Xiao, and Z. Y. Zhou, "ToF and stereo data fusion using dynamic search range stereo matching," *IEEE Transactions on Multimedia*, vol. 24, pp. 2739-2751, 2021, doi: 10.1109/tmm.2021.3087017.

[64] M. Poggi, G. Agresti, F. Tosi, P. Zanuttigh and S. Mattoccia, "Confidence Estimation for ToF and Stereo Sensors and Its Application to Depth Data Fusion," *IEEE Sensors Journal*, vol. 20, no. 3, pp. 1411-1421, 2020, doi: 10.1109/JSEN.2019.2946591.

[65] H. M. Huang, G. H. Liu, K. R. Duan, and J. Y. Yuan, "A convenient 3D reconstruction model based on parallel-axis structured light system," *Journal of Laser Applications*, vol. 33, no. 4, 82021, doi: 10.2351/7.0000469.

[66] Y. B. He, T. L. Yang, G. F. Lu, and J. G. Liao, "Study of Target Pose Measurement Based on Binocular Vision of Linear Structure Light," *Journal of Mechanical Engineering*, no. 5, pp. 1-3, 2017.

[67] F. Bruno, G. Bianco, M. Muzzupappa, S. Barone, and A. V. Razionale, "Experimentation of structured light and stereo vision for underwater 3D reconstruction," *ISPRS Journal of Photogrammetry and Remote Sensing*, no .66, pp. 508-518, 2011.

[68] C. L. Wen, J. B. Tan, F. S. Li, C. R. Wu, Y. T. Lin, Z. Y. Wang, and C. Wang, "Cooperative indoor 3D mapping and modeling using LiDAR data," *Information Sciences*, vol. 574, pp. 192-209, 2021, doi: 10.1016/j.ins.2021.06.006.

[69] Y. Deng, J. Xiao and S. Z. Zhou, "ToF and stereo data fusion using dynamic search range stereo matching," *IEEE Transactions on Multimedia*, vol. 24, pp. 2739-2751, 2022, doi: 10.1109/TMM.2021.3087017.

[70] F. Li, Q. L. Li, T. J. Zhang, Y. Niu, and G. M. Shi, "Depth acquisition with the combination of structured light and deep learning stereo matching," *Signal Processing: Image Communication*, vol. 75, pp. 111-117, 2019, doi: 10.1016/j.image.2019.04.001.

[71] X. X. Long, and X. J. Cheng, "三维视觉前沿进展 [Frontier Advances in 3D Vision]," *Journal of image and Graphics*, vol. 26, no. 6, pp.1389-1412, 2021.

[72] Z. Kang, J. Yang, Z. Yang, and S. Cheng, "A review of techniques for 3D reconstruction of indoor environments," *ISPRS International Journal of Geo-Information*, vol. 9, no. 5, pp. 339, 2020, doi: 10.3390/ijgi9050330.

[73] M. Vlaminck, H. Luong, W. Goeman, W. Philips, "3D scene reconstruction using omnidirectional vision and LiDAR: A hybrid approach," *Sensors*, vol. 16, no. 11. pp. 1923, 2016, doi: 10.3390/s16111923.

[74] G. Bianco, A. Gallo, F. Bruno, and M. Muzzupappa, "A comparative analysis between active and passive techniques for underwater 3D reconstruction of close-range objects," *Sensors*. vol. 13. No. 8, 2013, pp. 11007-11031, doi: 10.3390/s130811007.

[75] D. Ponsa, A. López, and F. Lumbreras, "3D vehicle sensor based on monocular vision," *In: Proceedings of the IEEE Intelligent Transportation Systems, NJ: IEEE Press*, pp. 1096-1101, 2005.