A Short Review on Vision-Based Object Grasping Automation with QR Code

Mooi Khee Teoh[1], Soo Fun Tan[2], Hou Pin Yoong[1#]
1 Biomechatronics Research Laboratory, Faculty of Engineering, Universiti Malaysia Sabah, Jalan UMS, 88400, Kota Kinabalu, Sabah, MALAYSIA.
2 Faculty of Computing and Informatics, Universiti Malaysia Sabah, Jalan UMS, 88400, Kota Kinabalu, Sabah, MALAYSIA.
Corresponding author. E-Mail: yoongpin@ums.edu.my; Tel: +6016-5210089; Fax: +6088- 320348.

ABSTRACT Rapid development in technologies has improved our quality of life at the same time application of automation into daily activities is becoming an imperative. The online order-picking system is foreseen to be a new normal of life and thus this paper reviews various past researches with related technologies. Few topics mainly on the vision-based grasping automation and monocular vision system using QR code as labels or markers is reviewed. The application of different types of grasping automation in variety field is studied and it shows that an Eye In Hand (EIH) type grasping automation, which the camera sensor is placed together with the robot arm's end, is suitable to be applied into an order-picking system. Thereafter, the monocular vision-based system is also reviewed. Studies found that monocular system is an effective method with low cost and easy installation process. Besides, monocular vision-based automation can operate at high accuracy and efficiency, with the aid of artificial markers such as QR code. QR code technology has been widely applied including products identification, item tracing, and manufacturing management. Additionally, QR codes can be used as markers for picking and packaging products in warehouse. However, limited research is observed using vision-based grasping automation system with QR code markers. Thus, a new research direction of monocular vision-based grasping automation using QR code is expected and suggestive.

KEYWORDS: Robotics; Object grasping; Machine vision; QR code; Artificial landmark; Pick-cart
Received 28 October 2020 **Revised** 03 December 2020 **Accepted** 13 December 2020 **Online** 2 December 2021

Review Article

INTRODUCTION

Fast growth in modern technologies allow automations to be applied into countless fields of work, not only in industrial environment but also in our daily lives. The development and application of automation has significantly improved our quality of life, at the same time, it minimizes manual operations done by human (Niu *et al*., 2018). Nowadays, daily living activities such as groceries purchasing can be done though online by just using an order-picking or the pick-cart system. Besides, with the strike of Covid-19 pandemic, people are now encouraged to do online shopping as it can effectively maintain a physical distance. Thus, the online pick-cart system is playing a crucial role at this moment and it might be a new normal to our life.

Various research had been carried out to design and implement a partly or fully automated order-picking system either by using human or robot as the picker (Jaghbeer *et al*., 2020). This review focus on the robot picker system where automation is used to carry out the picking operation. The robot picker is generally guided using sensors techniques. The sensors can be categorised into contacting and contactless types (Fantoni *et al*., 2014). For example, electrical sensor and mechanical switch are contacting type sensor while vision-based sensor such as a camera is a non-contacting type sensor. The sensor can be mounted on the robot or externally away from the robot to monitor the grasping process (Fantoni *et al*., 2014; Ishak & Mahmood, 2019). Vision-based automation that could reach and grasp a three-dimensional (3D) object is the highlight in this review. This is because visual system is an efficient and accurate solution with simple installation process and it is cost effective (Ramnath, 2004; Saxena *et al*., 2010). Besides, researches had shown that vision-based automation is capable to operate and function accurately without manual control from human (Bone *et al*., 2008; Zhihong *et al*., 2017).

Visual system can be categorised into mono and stereo system. Stereo or multi-cameras system uses two or more optical-based sensors while monocular vision system uses only one optical-based sensor. Therefore, monocular system saves the cost of camera and it can also reduce the calibration processes required as compared with stereo system. The major drawback of monocular system is lacking of depth information. However, object detection and positioning using monocular system can still be achieved with the aid of artificial labels or markers (Liu *et al*., 2020). An artificial landmark is a man-made label or marker which it can be designed with specific geometric shapes and having recognizable images. Although the system using artificial landmark consumes more time and effort to design and set up, yet it can be much efficient and accurate (Lin & Chen, 2011; Zhong *et al*., 2017; Liu *et al*., 2020).

In this review paper, some recent vision-based object grasping system and research related to vision-based automation that using artificial marker proposed by various researchers are reviewed and discussed. The organisation of the paper is as follows. Section 2 presents the application of object grasping system in recent years. Thereafter, Section 3 presents the monocular vision system with the aid of artificial landmark for autonomous system. Discussion and opportunities for monocular vision-based object grasping system with 2D marker are then presented in Section 4, and, finally, the conclusion is presented in Section 5.

OBJECT GRASPING SYSTEM

The employment of object grasping system plays a crucial role in variety fields such as food industrial, manufacturing, medical, and e-commerce (Fantoni *et al*., 2014; Pujari & Deosarkar, 2017; Chen & Chen, 2018). One of the common applications is in order-picking system where grasping automation is being used to pick and place the object from storage location to packing station (Jaghbeer *et al*., 2020; Liang *et al*., 2015). According to Liang *et al*. (2015) and Krug *et al*. (2016) the order-picking system with grasping automation has improved accuracy and reduced dependency on human labour. In addition, the introduction of grasping automation into the system enables the workplace to operate 24 hours with maximum functionality.

According to Ramnath (2004), Saxena *et al*. (2010) and Du *et al*. (2021) the key steps while performing object grasping are: 1) Detection and localization of target object; 2) Determination of grasping position and orientation; 3) Path planning and execution of grasping. Recent research had been focused on improving the grasping performance with better picking accuracy and shorter operating sequence (Ong & Joseph, 2014; Choi *et al*., 2018). Different sensing and control methods had been implemented to achieve the grasping principles either by geometrically driven or data-driven methods. Vision-based is preferable and recommended as it is cost effective, reliable and can be easily installed (Ramnath, 2004; Ishak & Mahmood, 2019). Besides, various research shown that vision-based grasping system able to perform accurate grasping operation even in cluttered environment (Bone *et al*., 2008; Saxena *et al*., 2010; Zhihong *et al*., 2017).

Fantoni *et al*. (2014) states that the installation of the camera can be either affixed on external place away from (Eye To Hand, ETH) or integrated together with the automation's end effector (Eye In Hand, EIH). The ETH types having wider field of view while EIH establish narrow view but clearer image. Thus, EIH is more suggested over ETH for order-picking system as EIH having higher accuracy, exact tracking positions and most importantly free mobility that able to function together with the automation to any place of the storage area (Levine *et al*., 2018; Ishak & Mahmood, 2019). There are total 9 researches referred in this study and 4 of them implementing ETH type, 4 researches using EIH method while only one research by Saxena *et al*. found to merge both types.

As shown in Figure 1, ETH type vision-based grasping automation is been used where the camera is installed on external place from the grasping automation (Ramnath, 2004; Krug *et al.*, 2016; Zhihong *et al*., 2017; Levine *et al*., 2018).

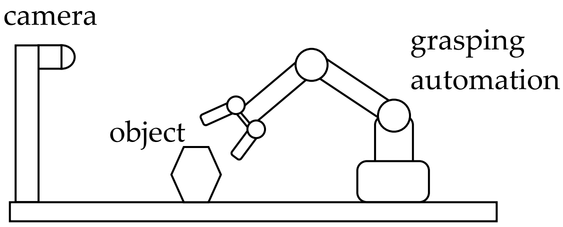

Figure 1. Eye To Hand (ETH) type vision-based grasping automation. (Ramnath, 2004; Krug *et al.*, 2016; Zhihong *et al*., 2017; Levine *et al*., 2018).

On the other hand, referring to Figure 2, Bone *et al*. (2008), Liang *et al*. (2015), Ishak & Mahmood (2019) and Yu *et al*. (2019) used EIH type grasping automation, which the camera is installed on the grasping automation.

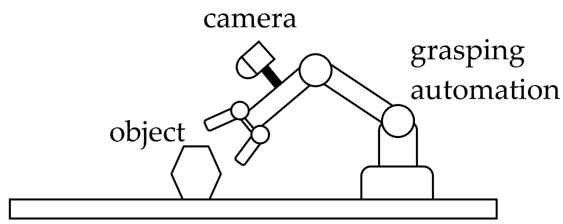

Figure 2. Eye In Hand (EIH) type vision-based grasping automation. (Bone *et al*., 2008; Liang *et al*., 2015; Ishak & Mahmood, 2019; Yu *et al*., 2019).

Moreover, in Saxena *et al*. (2010) studies, camera is set up on the robot (EIH method) to detect object's grasping point, while ETH types stereovision is for path planning and obstacles avoidance purpose as shown in Figure 3.

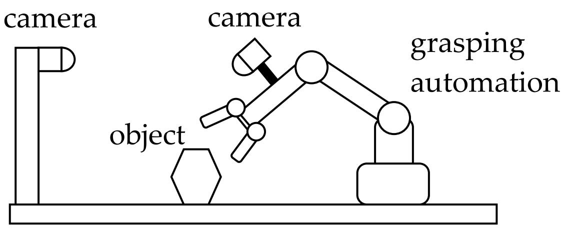

Figure 3. Eye To Hand (ETH) merge with Eye In Hand (EIH) type vision-based grasping automation (Saxena *et al*., 2010).

MONOCULAR VISION-BASED SYSTEM

As mentioned earlier, vision-based system is preferable and recommended for object grasping automation used by order-picking system. A vision-based system can be easily applied and installed, especially for indoor grasping automation (Zhihong *et al*., 2017; Yu *et al*., 2019). A system that uses only one optical sensor camera as the sensing input is known as monocular vision system. A monocular system with only one optical point lacks of depth information between the object and automation. However, the depth information can be calculated by taking at least 2 measurements from different optical points (Sazdovski & Silson, 2011; He *et al*., 2017).

Referring to Figure 4, parameter *f* is the focal length of camera, O_1 and O_2 is the object placed at different distance, d_1 and d_2 respectively, and I_1 and I_2 is image size of object captured by the camera, for the depth calculation. According to Cao *et al*. (2013) and He *et al*. (2017), the depth calculation of the object can be obtained based on basic pinhole imaging principle and Similar Triangles rules. The relationship between the parameters are shown in (Equations 1) and (Equation 2) while depth calculation is computed as (Equation 3).

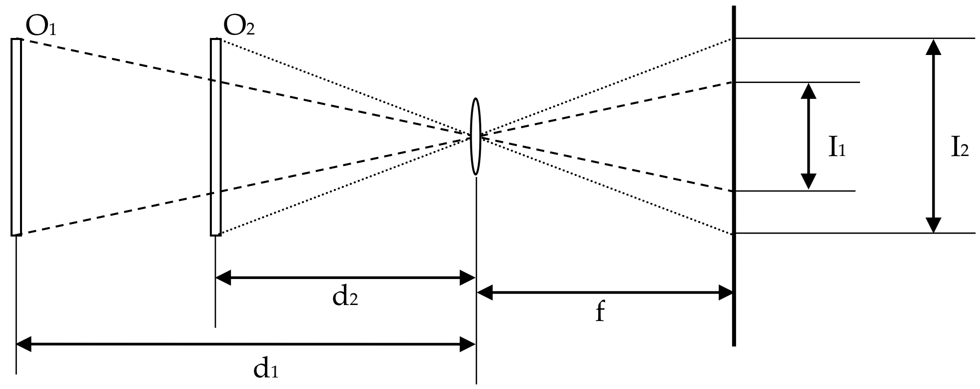

Figure 4. Depth calculation of monocular vision system (Cao *et al*., 2013).

According to Cao *et al*. (2013), the relationship of the object size (O_1, O_2), camera focal length (*f*), image size of object (I_1, I_2) and distance from the object to camera (d_1, d_2) can be computed using the following equations.

$$d_1 = \frac{f}{I_1} O_1 \tag{1}$$

$$d_2 = \frac{f}{I_2} O_2 \tag{2}$$

According to He *et al*. (2017), the depth information can be obtained from at least two images that captured the same object from different distance away. The image size of same object can be calculated using minimum two pairs of relative points of the same object which is:

$$d_2 = \frac{I_2}{I_1 - I_2} d_1 \tag{3}$$

On the other hand, a monocular system can be a great source of environmental information, but the extraction of required position information of specific object from the image acquired is challenging. Many previous researches use vision system with 2D artificial landmark to simplify the location features extraction process (Zhong *et al*., 2017; Yu *et al*., 2019; Liu, *et al*., 2020). In 2019, a fully automated indoor library system is developed with the aid of QR code as markers for object identification purpose (Yu *et al*., 2019). Besides, Zhong *et al*. (2017), Atali *et al*. (2018) and Lin & Chen (2011) uses artificial landmarks to automate the localizing and path planning process for indoor automation. The landmark can be placed on the walls, ceilings, floors as well as on the object itself. The use of landmark can reduce the calculation work and improve the recognition precision and detection process (Yu *et al*., 2019). According to Zhong *et al*. (2017), artificial landmarks in a given environment should be detectable, identifiable, and created easily.

With the increase resolution of imaging sensor, visual printed markers such as 1D and 2D barcode is a good option to be used as artificial markers. As compared with 1D barcodes, 2D barcodes have more flexible usage as it can store more information. 2D barcodes also easier to encrypt and can encode images, which can be read on almost any device with scanning functions such as smartphones (Cavanini *et al*., 2017). Moreover, 2D barcodes also feature exceptional fault tolerance, in which the information is still able to be decoded even if part of the code is damaged (Lin & Chen, 2011). There are few numbers of 2D barcodes types available in the market, these are such as the Quick Response code (QR code), PDF417, AZTEC, and Data Matrix code (Tan, 2020). Among the 2D barcode types, QR code is one of most common option to be used as markers and easy to be detected in indoors environment. Many researches used QR codes as marker as it is easy to produce, cost effective, flexible in size, and it can be modified whenever necessary (Atali *et al*., 2018; Cavanini *et al*., 2017; Yu *et al*., 2019). Figure 5 shows a sample of QR code. A QR code consists of Finder Pattern, a pattern to detect the position of the QR code as shown in the figure (Tan, 2020).

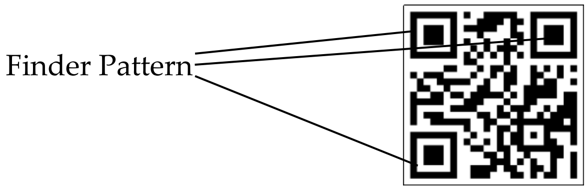

Figure 5. Example of 2D barcode types: QR Code.

The three recognisable points of finder pattern located at the corners of the QR code enables the monocular system to capture information of multiple points in a single image. Therefore, the depth information calculation can be done with only one image taken from an optical point. Besides, QR codes are designed with recognisable geometric shape that aids the detection of object during image processing. Thus, the drawbacks of a monocular system which are lacks of depth information and complicated process of location features extraction in image processing can be solved by introducing 2D QR code markers into the working environment.

MONOCULAR VISION-BASED OBJECT GRASPING SYSTEM WITH QR CODE MARKER

An order-picking system that using robot picker requires fast operation speed and high accuracy to fulfil the rapid order requirement. Review done shows that EIH grasping automation can perform accurately and able to operate automatically to handle parts and products with the aid of barcode marker (Pujari & Deosarkar, 2017; Chen & Chen, 2018; Yu *et al*., 2019). So, an EIH grasping types automation can be used as the robot picker either to move around at storage areas to pick up the items or to carry out the picking and packing at a picking station (Jaghbeer *et al*., 2020).

Moreover, studies on previous research also shown that marker-based monocular vision system that ideal for indoor working environment is both cost-effective and accurate. It is suitable for an order-picking system that most of the operations are done at indoor warehouse. However, to date, many researches using vision-based grasping automation with markers are observed (Pujari & Deosarkar, 2017; Chen & Chen, 2018), but very limited number of researches using QR codes as marker is observed (Yu *et al*., 2019). To the best of author knowledge to date, only one research is found using vision-based robot gripping system, and QR marker is used for an automated library system. In the system developed by Yu *et al*. (2019), binocular vision system had been utilized to capture the marker position. The use of monocular vision system with QR marker for a grasping automation is left unexplored. Therefore, combination of monocular vision and QR code technologies in the grasping automation of a fully automated order-picking system is foreseen to be a new research direction.

CONCLUSION

As a conclusion, the recent technology on grasping automation, monocular vison-based sensing system and system with QR code as marker had been study in detail and reviewed. The grasping automation can be applied in variety of field and EIH grasping types is suitable for operation with high precision as compared with ETH types. The review also found the monocular vision system is a better option than stereo system for its lower cost. Besides, the major drawback of monocular system can also be offset with the use of artificial marker. On top of that, QR code is a good choice among other types of barcode to be used as marker. This is because QR code not only used for positioning purpose, it also can be used for indicating information of the object. In short, a potential research direction is defined by combining the monocular vision-based sensing system, EIH type grasping automation and QR markers for implementation in an online order-picking system.

ACKNOWLEDGEMENTS

This is a part of a project supported by Universiti Malaysia Sabah Special Funding Scheme (SDK0216-2020).

REFERENCES

[1] Atali, G., Garip, Z., Ozkan, S. S. & Karayel, D. 2018. Path Planning of Mobile Robots Based on QR Code. *6th International Symposium on Innovative Technologies in Engineering and Science (ISITES).* 9-11 November, 2018, Antalya, Turkey. pp 31-38.

[2] Bone, G. M., Lambert, A. & Edwards, M. 2008. Automated Modeling and Robotic Grasping of Unknown Three-Dimensional Objects. *2008 IEEE International Conference on Robotics and Automation (ICRA)*. 19-23 May, 2008, California, USA. pp 292-298.

[3] Cao, Y.-T., Wang, J.-M., Sun, Y.-K. & Duan, X.-J. 2013. Circle marker based distance measurement using a single camera. *Lecture Notes on Software Engineering,* 1(4), 376-380.

[4] Cavanini, L., Cimini, G., Ferracuti, F., Freddi, A., Ippoliti, G., Monteriu, A. & Verdini, F. 2017. A QR-Code Localization System for Mobile Robots: Application to Smart Wheelchairs. *2017 European Conference on Mobile Robots (ECMR).* 6-8 September, 2017, Parism France. pp 1-6.

[5] Chen, Z.-Y. & Chen, C.-T. 2018. A remote controlled robotic arm that reads barcodes and handles products. *Inventions,* 3(1), 17-28.

[6] Choi, C., Schwarting, W., DelPreto, J. & Rus, D. 2018. Learning object grasping for soft robot hands. *IEEE Robotics and Automation Letters,* 3(3), 2370-2377.

[7] Du, G., Wang, K., Lian, S. & Zhao, K., 2021. Vision-based robotic grasping from object localization, object pose estimation to grasp estimation for parallel grippers: A review. *Artificial Intelligence Review,* 54(7), 1677-1734.

[8] Fantoni, G., Santochi, M., Dini, G., Tracht, K,. Scholz-Reiter, B., Fleischer, J., Lien, T. K., Seliger, G., Reinhart, G. & Franke, J. 2014. Grasping devices and methods in automated production processes. *The International Academy for Production Engineering (CIRP Annals),* 63(2), 679-701.

[9] He, L., Yang, J., Kong, B. & Wang, C. 2017. An automatic measurement method for absolute depth of objects in two monocular images based on sift feature. *Applied Sciences,* 7(6), 517.

[10] Ishak, A. J. & Mahmood, S. N. 2019. Eye in hand robot arm based automated object grasping system. *Periodicals of Engineering and Natural Sciences*, 7(2), 555-566.

[11] Jaghbeer, Y., Hanson, R. & Johansson, M. I. 2020. Automated order picking systems and the links between design and performance: A systematic literature review. *International Journal of Production Research,* 58(15), 4489-4505.

[12] Krug, R., Stoyanov, T., Tincani,V., Andreasson, H., Mosberger, R., Fantoni, G. & Lilienthal A. J. 2016. The next step in robot commissioning: Autonomous picking and palletizing. *IEEE Robotics and Automation Letters,* 1(1) , 546-553.

[13] Levine, S., Pastor, P., Krizhevsky, A. & Quillen, D. 2018. Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection. *The International Journal of Robotics Research,* 37(4-5), 421-436.

[14] Liang, C., Chee, K.J., Zou, Y., Zhu, H., Causo, A., Vidas, S., Teng, T., Chen, I., Low, K. & Cheah, C. 2015. Automated Robot Picking System for E-Commerce Fulfillment Warehouse Application. 2015. *Proceedings of the 14th International Federation for the Promotion of Mechanism and Machine Science (IFToMM) World Congress.* 30 October, 2015, Taipei, Taiwan. pp 398-403.

[15] Lin, G. & Chen, X. 2011. A robot indoor position and orientationm based on 2D barcode landmark. *Journal Of Computers (JCP)*, 6(6), 1191-1197.

[16] Liu, F., Zhang, J., Wang, J. & Li, B. 2020. The indoor localization of a mobile platform based on monocular vision and coding images. *ISPRS International Journal of Geo-Information (IJGI),* 9(2), 122.

[17] Niu, Z., Zhao Y., Hu, J., Wang, D., Dong, L. & Wang, L. 2018. Explore the development and application of electrical automation control systems. *Smart Construction Research,* 2(4), 1-6.

[18] Ong, J. O. & Joseph, D. T. 2014. A review of order picking improvement methods. *J@TI Undip: Jurnal Teknik Industri,* 9(3), 135-138.

[19] Pujari, T. S. & Deosarkar, S. 2017. Design of Intelligent and Robotic Library System. *2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT).* 19-20 May, 2017, Bangalore, India. pp 1903-1908.

[20] Ramnath, K. 2004. *A Framework for Robotic Vision-Based Grasping Task.* Project report, The Robotics Institute, Carnegie Mellon University.

[21] Saxena A., Wong L., Quigley M. & Ng A.Y. 2010. A vision-based system for grasping novel objects in cluttered environments. *In:* Kaneko M., Nakamura Y. (Eds.). *Robotics Research. Springer Tracts in Advanced Robotics*, (Vol 66). Berlin, Heidelberg, Springer.

[22] Sazdovski, V. & Silson, P. M. 2011. Inertial navigation aided by vision-based simultaneous localization and mapping. *IEEE Sensors Journal,* 11(8), 1646-1656.

[23] Tan, S.J. *QR code.* (https://foxdesignsstudio.com/uploads/pdf/Three_QR_Code.pdf). Last accessed on 15 October 2020.

[24] Yu, X., Fan, Z., Wan, H., He, Y., Du, J., Li, N., Yuan, Z. & Xiao, G. 2019. Positioning, navigation, and book accessing/returning in an autonomous library robot using integrated binocular vision and QR code identification systems. *Sensors,* 19(4), 783.

[25] Zhihong, C., Hebin, Z., Yanbo, W., Binyan, L. & Yu, L. 2017. A Vision-Based Robotic Grasping System Using Deep Learning for Garbage Sorting. *2017 36th Chinese Control Conference (CCC).* 26-28 July, 2017, Dalian, China. pp 11223-11226.

[26] Zhong, X., Zhou, Y. & Liu, H. 2017. Design and recognition of artificial landmarks for reliable indoor self-localization of mobile robots. *International Journal of Advanced Robotic Systems,* 14(1), 1729881417693489.