Picking Assistance System with MS-KINECT and Projection Mapping

Yuki Takehara, Yoshitoshi Murata
Faculty of Software and Information Science
Iwate Prefectural University
Takizawa, Japan
e-mail: g031m104@s.iwate-pu.ac.jp, y-murata@iwate-pu.ac.jp

Takamichi Yoshikawa
Aioi Systems Co., Ltd
Ota-ku, Tokyo, Japan
e-mail: yoshikawa@hello-aioi.com

***Abstract*—Industry 4.0 is a well-regarded concept for automation in manufacturing. However, a shortage of high-skilled workers has necessitated realistic solutions for establishing high productivity and quality. We propose an information and communications technology (ICT) picking assistance system to lower human errors for high quality. In this system, MS-KINECT detects whether a hand is inserted into the correct cell of a shelf to pick items and whether a hand is inserted into the correct box in a cart to put items in. The miss-detection rate for wrong operation in the picking process is very low in an experimental evaluation, and we expect it to be very close to zero in a near future. We determine that the proposed picking error detection function would be useful for business. However, we must improve the delivery detection accuracy because of its higher detection error rate in an experimental evaluation. In this system, projection mapping technologies are used to indicate which cell items should be picked from, instead of using a lamp. The indicating system which uses a projector lowers the introduction cost than existing one when compared to the existing system using a lamp. We clarified that gray sandpaper is one of the best materials to serve as a tag for MS-KINECT to recognize indicated colors and digits.**

Keywords-Smart factory; Industry 4.0; picking; KINECT; projection mapping.

I. INTRODUCTION

The German government's Industry 4.0 (ID4) initiative *Industrie 4.0* has revolutionized Germany's manufacturing industry [1], with goods in "smart factories" being moved, picked, and delivered automatically [2]. ID4 technologies are useful for establishing high productivity and quality in light of a shortage of high-skilled workers. If the latest manufacturing robotics and custom assembly lines were introduced in model factories for ID4, products would be automatically conveyed and assembled, and there would be fewer workers. However, it is impossible for most existing factories to replace all of their manufacturing lines with more advanced ones. Realistic solutions for establishing high productivity and quality in light of a shortage of high-skilled workers are as follows:

(1) Developing and introducing information and communications technology (ICT) systems to assist a low-skill worker to be close to a high-skill worker and lower the human error.
(2) Replacing workers with robotics step by step.

In this paper, we propose an ICT picking assistance system to lower human errors in the picking process. The picking process is when a worker picks items, such as assembly parts out of numbered cells on shelves and puts them into boxes. For example, in an automobile assembly factory, a worker takes different parts from cells of a shelf and puts them into assortment boxes corresponding to production orders. The parts in each cell are the same, and different items are stored in different cells. Assortment boxes are carried to workers on an assembly line. If a worker picks parts from an incorrect cell and the parts are subsequently assembled into a product, it would take too much time to detect the incorrect parts and exchange them with correct ones. In the worst case, an automobile assembled with incorrect parts could be shipped. These processes in which such mistakes can occur are widely used, not only in assembly factories, but also at delivery centers.

Even though picking operations are monotonous, completely preventing mistakes is difficult because workers are human. Therefore, there are several kinds of picking assistance systems for decreasing incorrect pickings. Aioi Systems Co. Ltd. to which one of co-authors belongs provides the digital picking system "L-PICK," which indicates the cells of a shelf and the number of items to pick from those cells by lighting a lamp mounted on each cell [3]. However, since L-PICK does not have a function that detects incorrect picking, completely preventing it is impossible. Many companies request Aioi Systems Co. Ltd. to develop and provide the picking error detection system to lower recovery cost. Hence, we developed an operation error detection system for the picking process with MS-KINECT [4]. In this system, two sets of MS-KINECT v1 [5] trace a hand wearing a colored glove from diagonally backward and monitor whether the hand is inserted into a cell. Since this system uses a color tracing function to measure the position of a hand, it cannot be applied to a food material delivery service because of differences in food color.

This time, we developed a new picking assistance system for the picking process. It uses a set of MS-KINECT v2 mounted on the top of a shelf. Since our partner Aioi Systems Co. Ltd. has developed a new cell lighting technology that uses a projector [3], we used it in our new system. Our system recognizes the position of a cell lit by a projector and the number presented on a tag, detects the position of a cell into which a hand is inserted, counts the number of times a hand has been inserted into the cell, and compares them with recognized ones. In addition to picking functions, we have developed delivery functions. Our delivery assistance function detects whether a hand that has items is inserted into the correct delivery box.

The miss-detection rate for wrong operation in the picking process is very low in an experimental evaluation, and we expect it to be very close to zero in a near future. We determine that the proposed picking error detection function would be useful for business. However, we must improve the delivery detection accuracy because of its higher detection error rate in an experimental evaluation.

After introducing related work in Section II, detection methods for incorrect operations are introduced in Section III. Detection technologies in which MS-KINECT are used are introduced in Section IV. Experiments and results are described in Section V. Conclusions and future work are described in Section VI.

II. RELATED WORK

Human beings have excellent abilities. Workers in an automobile assembly factory can use their sense of vision and touch to detect subtle depressions or distortions that a computer system cannot. On the other hand, human beings sometimes make mistakes. Several kinds of assistance systems that decrease the number of mistakes have therefore been developed. Existing picking assistance systems are introduced in this section.

A picking assistance system has some of the following four functions:

(1) Indicating a cell of a shelf from which an item should be picked and the number of items that should be picked from that cell.

(2) Detecting whether a worker has picked the correct number of items from the correct cell.

(3) Indicating the box or cell of a tray in which picked items should be stored.

(4) Detecting whether a worker has put items in the correct box or cell of a tray.

There are several kinds of system for indicating cells. Aioi Systems Co. Ltd. provides the digital picking system "L-PICK," which indicates cells of a shelf and the number of items to pick from those cells by lighting a lamp mounted on each cell [3]. There are several systems in which a Head-Up Display (HUD) and augmented reality technology are used to assist picking operations. Schwerdtfeger used a semi-transparent HUD and augmented reality technology enabling a worker to see an arrow or frame displayed in front of a cell of a shelf [6]. Baumann used a single-eye HUD, and a worker recognized a cell from which he or she would pick up items with guidance displayed on a mirror of the HUD [7][8]. Guocompared the HUD, cart-mounted display (CMD), Light, and Paper Pick List as picking assistance systems [9]. This system also provided the delivery assistance function. They concluded that the pick-by-HUD and pick-by-CMD were superior by all metrics than the current pick-by-paper and pick-by-light systems, but the differences between the HUD and CMD were not significant and did not show that a HUD was better than a CMD. However, experimental results should be different in other experimental conditions. In practical situations, the number of cells (in that paper they were called bins) is usually less than 12, and a worker can see a lighted lamp at a glance. And multiple lamps are not lighted simultaneously; a single lamp is lighted for each occurrence where an item is picked up. Therefore, the practical error rates and task times recorded in this study would produce better results than those in that study. Furthermore, it is not certain whether a worker should have a palm-size PC and wear a HUD for extended periods. In our research, most workers would not like to carry a barcode reader.

As described above, systems indicating a cell by a lamp have been used in business. Likewise, indicating systems using an HUD have been developed in research.

We consider a picking and delivery error detection method in the next section.

III. PICKING AND DELIVERY ERROR DETECTION METHODS

In this section, prospective methods for detecting whether items are picked from the correct cell and delivered to the correct box are introduced and evaluated. This time, in addition to detecting when items are picked from a cell, these methods determine when a hand is inserted into a cell, when a tag, such as the barcode attached to a cell is read, and when a picked item is dropped, and a new item is picked. The following are the prospective methods:

(1) Reading a **barcode** attached on a cell with a barcode reader.

(2) Reading a **passive Radio Frequency (RF)-ID** set on a cell with a RF-ID reader.

(3) Reading an **active RF-ID** set on a cell with a RF-ID reader.

(4) Detecting change of weight with a **load sensor**.

(5) Detecting when a hand and/or arm is inserted into the correct cell with a **photoelectric sensor**.

(6) Detecting when a hand and/or arm is inserted into the correct cell with a **3D camera,** such as MS-KINECT.

The above methods are narrowed down by the following evaluation criteria:

(1) Additional cost to introduce a detection function.

(2) Additional operations for a worker.

(3) Detection accuracy.

An evaluation of the picking error detection methods is shown in Table I. As for barcodes, the cost of attaching a barcode to each cell is cheap, and barcode readers are not expensive. However, carrying a barcode reader and scanning barcodes are cumbersome for workers.

Passive RF-ID presents the same difficulties as barcodes. In addition to having to carry a RF-ID reader, the weak signal strength of passive RF-ID requires positioning the reader in close proximity to a RF-ID tag.

As for active RF-ID, despite having to carry the reader, it does not need to be positioned in close proximity to a RF-ID tag because the signal strength is strong. However, because of their strong signal strength, active RF-ID readers sometimes read RF-ID tags placed in other cells.

As for load sensors, their detection accuracy is high. However, they are usually expensive, and each sensor must be wired to a PC. Introduction costs are therefore high. The same holds true for photoelectric sensors.

3D cameras using MS-KINECT usually cost a few hundred dollars. While introduction costs would be high under our proposed system because one MS-KINECT set would be required per shelf, our system alleviates the need for workers to carry a reader, and the detection accuracy is high. We have determined that 3D cameras would be the best method overall for our picking assistance system.

The following three prospective methods are considered for delivery:

(1) Detecting change of weight with a **load sensor**.
(2) Detecting whether a hand is inserted into the correct box with a **photoelectric sensor**.
(3) Detecting whether a hand is inserted into the correct box with a **3D camera,** such as MS-KINECT.

The evaluation criteria for delivery is the same as those for picking. The evaluations for the above three methods are the same as those in Table I. We think a method using a 3D camera is the best for delivery when its accuracy is high.

TABLE I. EVALUATION OF PICKING ERROR DETECTION METHODS

Method	A. Cost	A. Operation	Accuracy
Barcode	Low	Big	Middle
Passive RF-ID	Middle	Big	Low
Active RF-ID	Middle	Little	Middle
Load	High	Little	High
Photoelectric	High	Little	Middle
3D camera	Low	Little	This paper

IV. DETECTION TECHNOLOGY

In this section, a detection technology by which a worker picks items from a cell, delivers them to a box, and a reading technology that detects the number on a tag lit by a projector are introduced.

A. *Detection technology for picking items from the correct cell*

As described in the previous section, a technology that detects whether a worker picks items from an indicated cell is needed. Since detecting whether a worker is picking items from a cell is difficult, we decided to focus on detecting when a hand is inserted into a cell and counting the number of times a hand is inserted into an indicated cell instead of detecting and counting hands picking items from an indicated cell. The motion monitoring function in MS-KINECT is popularly used to estimate position of joints on the body. However, the MS-KINECT must be set in front of the body between 0.5 – 5 m. It is impossible to set the MS-KINECT in front of a worker in a factory or delivery center. Therefore, Open-CV color tracing technologies [10] were used to trace hands in our previous system. And, since two MS-KINECT sets were needed for each shelf, the system was not economical. Therefore, we decided not to use the color tracing technology and opted to detect a hand inserted into a cell using a MS-KINECT set.

From past experiments we have determined that the best mount position for a MS-KINECT to detect whether a hand enters a cell is just above the surface of a shelf aperture. The MS-KINECT 3D camera searches for a hand and arm just over the surface of a shelf aperture as shown in Fig. 1. The MS-KINECT must be set at a position in which its 3D camera can observe the entire shelf aperture. This system detects whether a hand is inserted by changing the depth in front of a cell. When a hand and/or arm is inserted into a cell, the depth in such a view is changed from L_f to L_h. A change in depth L_h corresponds to the length between the MS-KINECT and the hand and/or arm. Its position is within the cell aperture in which the hand and/or arm is inserted.

The coordinates of the four corners of each cell are pre-set before estimating the cell number. In Fig. 1, the coordinates of 16 corners are pre-set. The number of the cell in which a hand is inserted is estimated by comparison between a coordinate of the detected hand and the coordinates of four corners for each cell ⓝ.

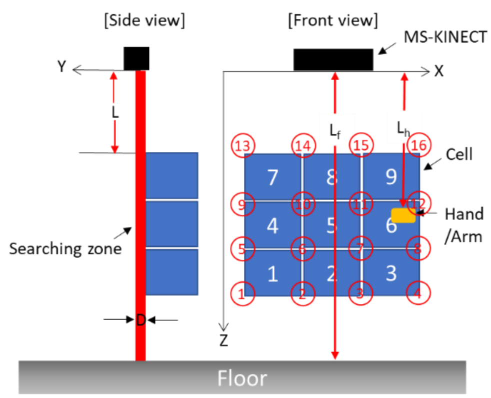

Figure 1. Mounting position of MS-KINECT and searching zone

B. *Lighting a tag and number*

Since our partner AIOI System Co. Ltd. has developed the new lighting method into which the projection mapping technology is used to indicate a picking cell, our experimental system uses this projection mapping technology. A very short focal projector mounted near the MS-KINECT lights a tag attached to a cell and projects a digit on it to indicate the cell and the number of items to be removed as shown in Fig. 2.

A computer knows which tag of a cell is lit, so there is no need for it to detect which tag is lit with the MS-KINECT. In this case, the MS-KINECT is connected to a computer. However, we plan to develop a picking robot that picks items

 ISBN: 978-1-61208-616-3

up and puts them into an indicated box in the near future. Since the robot must detect which tag is lit and read a digit on it, we developed a technology that realizes the above functions with the MS- KINECT. In this system, the font used for digits is a seven-segment font as shown on the right side of Fig. 2. Our system recognizes which kind of number is presented by detecting which segments are white.

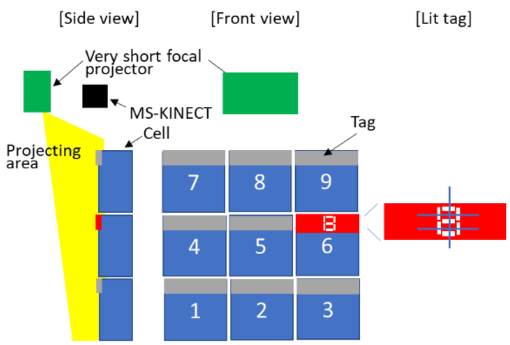

Figure 2. Layout of projector and cells

C. *Detection technology for delivering items*

We imagine a delivery cart as shown in Fig. 3. A MS-KINECT is mounted to search the surface of a mass of boxes. The search zone is just above the mass of boxes. When a hand/arm or item is inserted into a box, the PC on this cart detects whether a hand/arm or item is inserted into the box by the change in depth.

The number of the box that a hand is inserted into is estimated to compare it with a coordinate of the detected hand and the coordinates of four corners for each box, the same as for a cell in Section IV-A.

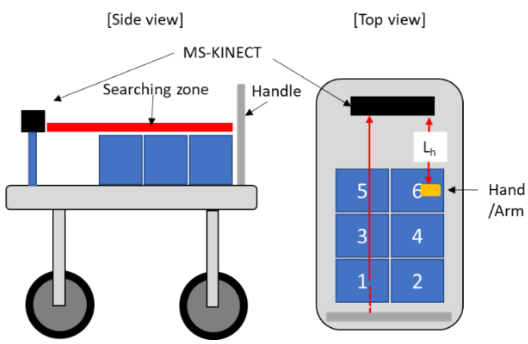

Figure 3. Image of delivery cart

V. EXPERIMENT

We developed an experimental shelf as shown in Fig. 4. An MS-KINECT is mounted 65.5 cm away from the shelf. The shelf consists of 3 x 3 cells. The size of the shelf is 67.5 x 64.5 cm, and the size of each cell is 22.5 x 21.5 cm. The length between the floor and the bottom of the shelf is 98 cm. We measured the error rates for detecting a hand inserted into a cell and whether the MS-KINECT can recognize a lit tag and the number on it using the experimental shelf.

Before estimation, the coordinates of the corners of each cell are pre-set using the pre-set windows shown in Fig. 5. The corner number is selected with the corner number button. The coordinates of each corner are entered by clicking a corner or inputting digits. The red grid of the shelf aperture is generated by clicking the grid button.

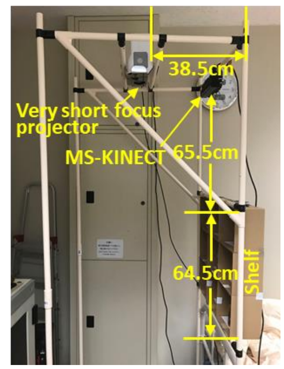

Figure 4. Experimental shelf

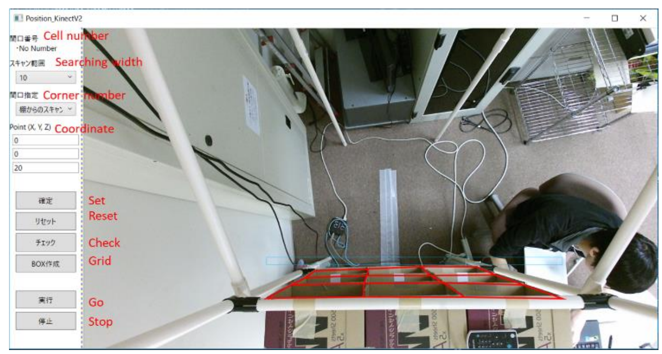

Figure 5. Pre-set window

A. *Detection error rate for picking operation*

The detection error rate for picking operation would change depending on the width D of the searching zone, the searching period, the cell position, and the threshold width to detect a hand and/or arm. We measured the detection error rate under conditions in which the width D of the searching zone is 1 cm / 3cm / 5 cm / 10 cm, the searching period is 500 msec. / 1000 msec., and the threshold width for detecting a hand and/or arm is 1 cm. Beer cans were used as picking items. The number of participants was ten. Each participant picks an item from each cell five times. The average detection error rates of every cell vs. the width of searching zone D are 1 cm / 3cm / 5 cm / 10 cm as shown in Fig. 6. The parameter of this figure is the searching period. The detection error rate in 500 msec. is lower than that in 1000 msec. in each the searching width D. The average error rate for each cell in which the searching period is 500 msec. is shown in Table II. The error rate increases in accordance with an increase in width D. This is because a participant tends to insert his/her hand into a cell through the searching zone in front of other cells.

 ISBN: 978-1-61208-616-3

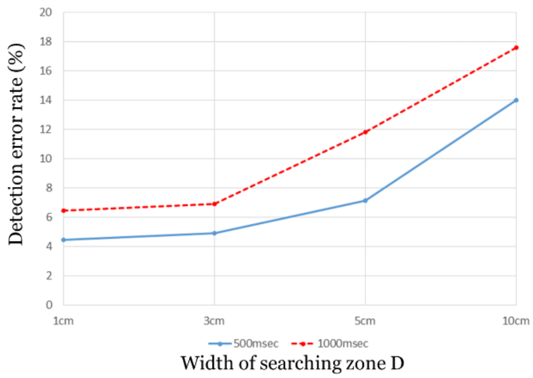

Figure 6. Average detection error rates for every cell

TABLE II. AVERAGE DETECTION ERROR RATES (%) FOR EACH CELL

Cell #	1cm	3cm	5cm	10cm
1	2	0	0	4
2	4	4	4	6
3	4	6	4	0
4	4	0	0	1
5	6	6	2	8
6	2	8	14	20
7	6	0	10	24
8	4	16	10	28
9	8	4	20	26

The most serious problem in the picking error detection system is that the system recognizes wrong operations to be fair. We measured the miss-detection rate for wrong operation. In this experiment, a correct cell is No. 5. The number of participants is ten. Each participant picks an item from cells around the No. 5 cell five times. The detection rate for wrong operations to be wrong, detection rate for wrong operation to be fair and practical detected wrong operations are shown in Table III. Average detection rate for wrong operations to be wrong is 95.7 %. And, the proposed system did not recognize wrong operations to be fair, completely detected wrong operations.

TABLE III. MISS-DETECTION RATE FOR WRONG OPERATIONS

Cell #	Detection Rate (%)	No. 5 cell D. Rate (%)	Detected errors
1	98	0	1-6
2	98	0	2-3
3	96	0	3-6, 6
4	98	0	1-4
5			
6	96	0	6-9, 3-6
7	94	0	(4-7) x 3
8	96	0	(5-8) x 2
9	90	0	(6-9) x 5

However, the system recognized that a participant picked an item from the No. 6 cell, even though he picked it from the No. 3 cell. The reason of this error detection is that the system detects an item in front and within 1 cm from the No. 6 cell after picking from the No. 3 cell. We think this miss-detection rate is low, but not enough. This reason would cause the detection for wrong operation to be fair. We plan to constitute a few cm non-detection area around each cell, and guard time not to detect after detecting an item to be picked from a cell. We think these constitutions would lower the detection rate for wrong operation to be fair exceedingly close to zero.

B. Recognition of a lit tag and the number on it

We noticed that the color through the video camera of the MS-KINECT was very different from the color we recognized and that the color through the video camera of the MS-KINECT changed in accordance with the color and luster of a tag. Example colors on a sheet of white paper, black paper, and gray sandpaper are shown in Fig. 7. The differences between the colors as displayed on a smartphone and those as displayed on the MS-KINECT are shown in Fig. 8. The colors displayed on a smartphone are almost equal to those seen with the naked eye. We selected red, green, and blue as the colors projected on a tag. The color characteristics of the MS- KINECT are very different from those of a smartphone. As a result, gray sandpaper is the best material for representing original colors. Our system can read every number perfectly on a red, green, or blue background. When implementing systems for clients, these three colors are usable.

Figure 7. Colors on tags as displayed on video camera of MS-KINECT

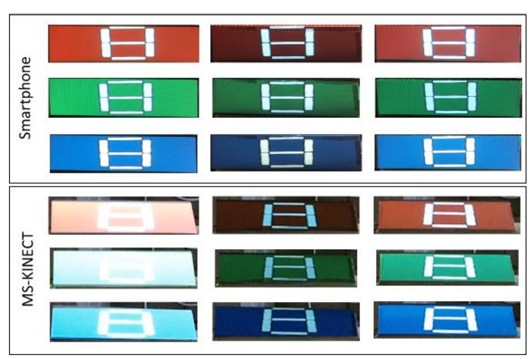

Figure 8. Colors as displayed on smartphone and MS-KINECT

C. Delivery error rate

We measured the delivery error rate using six boxes on a table as shown in Fig. 9 instead of using a delivery cart. The MS-KINECT is placed on another table. It was placed 60 cm from the top of the boxes. From the results of Experiment A, we decided that the searching period is 500 msec. and that the width of the search zone is 1 cm.

Since we noticed that our system easily detected multiple boxes, we constituted a 5 cm wide non-detection area on boxes that are on the near side of a worker, as shown in Fig. 9. When the depth far from the boxes was not fixed, the detection accuracy was poor and unstable. We set the screen low to fix the maximum depth from the MS-KINECT. The average error rate for each box is shown in Table IV. The number of participants was five, the same as in the picking experiment. Each participant puts an item into each box five times. Overall, the error rate, especially the double count rate, is high.

In some cases, a worker might throw an item into a box, so we also measured the delivery error rate when items were thrown into a box. However, our system could not detect a thrown item because the searching period of 500 msec. was longer than the necessary period for detecting an item.

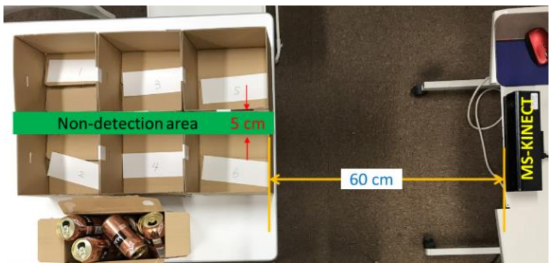

Figure 9. Experimental system for measuring delivery error rate

TABLE IV. AVERAGE DELIVERY ERROR RATES (%) FOR EACH BOX

Box number	No-detection	Double count	Total
1	0	12	12
2	0	8	8
3	8	20	28
4	4	0	4
5	24	4	28
6	0	12	12

VI. CONCLUSION AND FUTURE WORK

A realistic solution for establishing high productivity and quality for the picking process in light of a shortage of high-skilled workers is to introduce a picking assistance system that detects incorrect operations by workers. We introduced a picking and delivery assistance system in which an MS-KINECT detects whether a hand is inserted into the correct cell of a shelf to pick items and whether a hand is inserted into the correct box on a cart to put items in.

The miss-detection rate for wrong operation is very low in this system, and it is possible to be exceedingly close to zero in a near future. We determine that the proposed picking error detection function would be useful for business. The other hand, we must improve the delivery detection accuracy. We try to adopt the MS-KINECT motion monitoring function to estimate the position of hand.

ACKNOWLEDGEMNT

Thanks to Kazuhiro Yoshida for help in performing this research. This project is partially supported with the competitive research funding by Iwate Prefectural University.

REFERENCES

[1] Industrie 4.0; http://www.bmwi.de/Redaktion/EN/Dossier/industrie-40.html, [retrieved: February, 2018].

[2] What is a smart factory? ; http://www.bmwi.de/Redaktion/EN/FAQ/Industrie-4-0/faq-industrie-4-0-03.html, [retrieved: February, 2018].

[3] AIOI SYSTEMS Co., LTD Digital Picking System Products Introduction; https://www.hello-aioi.com/en/product/product-list/, [retrieved: February, 2018].

[4] Y. Uda, K. Yoshida, and Y. Murata, "Operation Error Detection System in Picking Processes by Camera Depth Sensors," IPSJ, Transaction of Consumer Device and Service, Vol.6 No.1, pp. 1-12, 2016, [written in Japanese].

[5] Meet Kinect for Windows; https://developer.microsoft.com/en-us/windows/kinect, [retrieved: February, 2018].

[6] B. Schwerdtfeger, and G. Klinker, "Supporting order picking with augmented reality," Proceedings of 7th IEEE/ACM International Symposium on Mixed and Augmented Reality, pp. 91–94, 2008.

[7] H. Iben, H. Baumann, C. Ruthenbeck, and T. Klug, "Visual based picking supported by context awareness: Comparing picking performance using paper-based lists versus list presented on a head mounted display with contextual support," Proceedings of 11th International Conference on Multimodal Interfaces, pp. 281–288, 2009.

[8] H. Baumann, T. Starner, H. Iben, and P. Zschaler, "Evaluation of graphical user-interfaces for order picking using head-mounted displays," Proceedings of 13th International Conference on Multimodal Interfaces, pp. 377–384, 2011.

[9] A. Guo et al., "A comparison of order picking assisted by head-up display (HUD), cart-mounted display (CMD), light, and paper pick list, "Proceedings of 2014 ACM International Symposium on Wearable Computers, pp. 71–78, 2014.

[10] Open-CV; https://opencv.org/, [retrieved: February, 2018].

 ISBN: 978-1-61208-616-3