

반도체 자동화를 위한 빈피킹 로봇의 비전 기반 캘리브레이션 방법에 관한 연구

구교문* · 김기현** · 김효영** · 심재홍**†

*한국공학대학교 IT반도체융합공학부, **한국공학대학교 메카트로닉스공학부

A Study on Vision-based Calibration Method for Bin Picking Robots for Semiconductor Automation

Kyo Mun Ku*, Ki Hyun Kim**, Hyo Yung Kim** and Jae Hong Shim**†

*Department of IT Semiconductor Engineering, Tech University of Korea,

**†Department of Mechatronics Engineering, Tech University of Korea

ABSTRACT

In many manufacturing settings, including the semiconductor industry, products are completed by producing and assembling various components. Sorting out from randomly mixed parts and classification operations takes a lot of time and labor. Recently, many efforts have been made to select and assemble correct parts from mixed parts using robots. Automating the sorting and classification of randomly mixed components is difficult since various objects and the positions and attitudes of robots and cameras in 3D space need to be known. Previously, only objects in specific positions were grasped by robots or people sorting items directly. To enable robots to pick up random objects in 3D space, bin picking technology is required. To realize bin picking technology, it is essential to understand the coordinate system information between the robot, the grasping target object, and the camera. Calibration work to understand the coordinate system information between them is necessary to grasp the object recognized by the camera. It is difficult to restore the depth value of 2D images when 3D restoration is performed, which is necessary for bin picking technology. In this paper, we propose to use depth information of RGB-D camera for Z value in rotation and movement conversion used in calibration. Proceed with camera calibration for accurate coordinate system conversion of objects in 2D images, and proceed with calibration of robot and camera. We proved the effectiveness of the proposed method through accuracy evaluations for camera calibration and calibration between robots and cameras.

Key Words : Bin picking technology, Calibration, RGB-D cameras, Vision-based calibration method

1. 서 론

반도체 뿐만 아니라 많은 산업 현장에는 부품을 다양하게 생산하고 조립하여 제품을 완성한다. 생산된 부품은

종종 여러 부품이 혼재되어 상자 안에 담기게 된다. 무작위로 담긴 부품들 중 필요 부품만 꺼내어 분류 정렬 작업을 진행하고 이때 많은 시간과 인력이 소요된다. 또한 후속 공정에서 필요한 부품을 올바르게 꺼내 정렬하고 전달하는데 필요하다.

3차원 공간에서 다양한 물체와 로봇, 카메라의 위치나 자세 등 서로 간의 정보를 알아야 가능하므로 무작위로

†E-mail: jhshim@tukorea.ac.kr

담긴 부품들의 분류 정렬의 자동화는 어렵다. 기존에는 특정 위치에 있는 물체만을 로봇이 잡거나 사람이 직접 물건을 분류하여 정렬하는데 이러한 분류 및 정렬 작업을 로봇이 특정 부분을 담당하지 않고 3차원 공간에서 임의의 물체를 잡기 위해서는 빈피킹(Bin Picking) 기술이 필요하다[14].

빈피킹 기술을 실현하기 위해 로봇과 파지 대상 물체 그리고 카메라 간의 좌표계 정보를 파악해야 한다[5-6]. 카메라에서 인식한 물체를 잡기 위해서는 서로의 좌표계 정보를 파악하는 캘리브레이션(Calibration) 작업이 반드시 필요하다.

카메라에서 보이는 2D 이미지에서 3차원에 해당하는 타겟(Charuco Board)의 자세를 추정하여도 깊이 값이 얇은 추정하기가 어렵다. 따라서 RGB-D 카메라의 Depth 정보를 이용하여 이동 정보 중에 Z값에 대입하여 빈피킹 로봇을 위한 비전 기반 캘리브레이션 방법에 관한 연구를 진행한다. 해당 방법이 카메라 캘리브레이션, 그리고 로봇과 카메라 간의 캘리브레이션에서 얼마나 정확한지 정확성 평가를 통해 입증한다.

2. 비전 기반 캘리브레이션

2.1 좌표계 정의

로봇과 카메라의 캘리브레이션은 서로 간의 좌표계 관계를 아는 것이다. 로봇의 베이스(Base)와 그리퍼(Gripper), 타겟, 그리고 카메라의 좌표계를 정의하고 이들 간의 관계를 정의한다.

2.1.1 베이스와 그리퍼의 좌표계 관계

베이스의 경우 Fig 1과 같이 X_b, Y_b, Z_b 로 좌표계의 각 축을 나타내고 그리퍼의 경우 로봇의 끝점을 의미하며 X_g, Y_g, Z_g 로 좌표계를 표현한다. 또한 둘 간의 관계인 베이스부터 그리퍼까지(gT_b) [7]는 주황색 화살표의 관계를 의미한다. 여기서 T 의 경우 식(1)과 같이 회전 성분인 r 과 이동 성분인 t 가 포함된 동차변환행렬식(Homogeneous Matrix)을 의미한다. 반대 방향의 관계(파란색) bT_g 는 gT_b 의 역행렬을 통해 구할 수 있다.

$$T = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_{11} \\ r_{21} & r_{22} & r_{23} & t_{21} \\ r_{31} & r_{32} & r_{33} & t_{31} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (1)$$

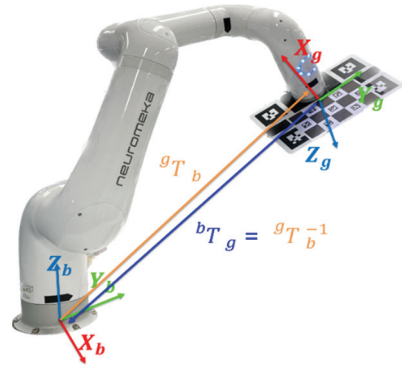


Fig. 1. Coordinate System from Base to Gripper.

2.1.2 카메라와 타겟의 좌표계 관계

카메라의 경우 Fig 2과 같이 X_c, Y_c, Z_c 로 좌표계의 각 축을 나타내고 타겟의 경우 X_t, Y_t, Z_t 로 좌표계를 표현한다. 또한 둘 간의 관계인 타겟부터 카메라까지(cT_t)는 주황색 화살표의 관계를 의미한다.

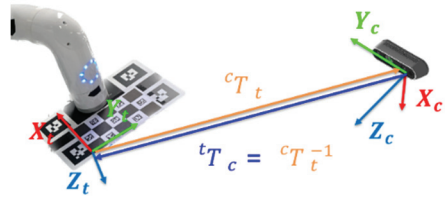


Fig. 2. Coordinate System from Target to Camera.

2.1.3 카메라 캘리브레이션

카메라 캘리브레이션은 정확한 이미지를 얻어 현실을 2D 이미지로 표현하기 위해 반드시 필요하다. 3D 좌표를 복원해야 하므로 정확한 이미지 변환은 매우 중요하다. 카메라 캘리브레이션은 Fig 3과 같이 실제 크기(체스 박스 40mm, Aruco 박스 20mm)를 알고 있는 물체의 이미지를 촬영하여 카메라의 내부 파라미터와 외부 파라미터를 얻

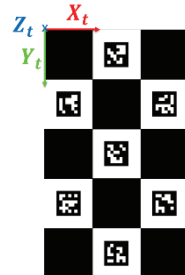


Fig. 3. Coordinate System Charuco Board.

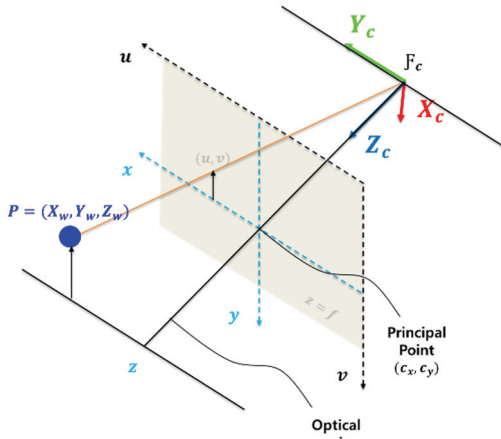


Fig. 4. Pin-Hole Camera Model.

어내는 것이다. [8]에 따르면 Fig 4은 핀홀(Pin-Hole) 카메라의 3D에서 2D 투영 모델이다. 실제 물체에 있는 점 P를 (u, v) 이미지 평면에 투영하여 2D 이미지를 나타낸다.

(u, v) 이미지 평면은 식(2)과 같이 카메라 내부 파라미터인 (3×3) 행렬과 (3×4) 행렬로 이루어진 세계 좌표계와 카메라 좌표계의 회전과 이동 변환 정보 그리고 세계 좌표계의 3D 점으로 표현한다.

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (2)$$

2.1.4 Eye to Hand 캘리브레이션

Eye to Hand 캘리브레이션은 카메라가 외부에서 고정되어 타겟을 관찰하여 로봇과 카메라 간 좌표계 관계를 알아내는 것이다. 로봇과 카메라 캘리브레이션에서 주로 다루어지는 것은 $AX=XB$ 문제이며 이 문제는 Eye to Hand의 경우 Fig 5와 같이 그리퍼와 타겟의 관계를 모를 때 사용한다. 이때 그리퍼와 타겟은 단단히 고정되어 자세가 바뀌어도 변하지 않아야 한다[9-13].

Fig 5에서 A 행렬은 첫 번째 자세(1)에 해당하고 B 행렬은 두 번째 자세(2)에 해당한다. 이렇게 서로 다른 위치와 방향으로 움직여 진행하고 이처럼 다수의 위치 사이의 X를 풀어 에러를 줄이는 방식으로 진행한다. 여기서 T는 동차변환행렬(4x4)이며 이 행렬 구성은 식(1)과 같다. 이 행렬은 회전행렬(3x3)과 이동 행렬(3x1)로 구성되어 있다 [14].

Eye to Hand 캘리브레이션에서 첫 번째 자세와 두 번째 자세를 연립하게 되면 수식(3)과 같다.

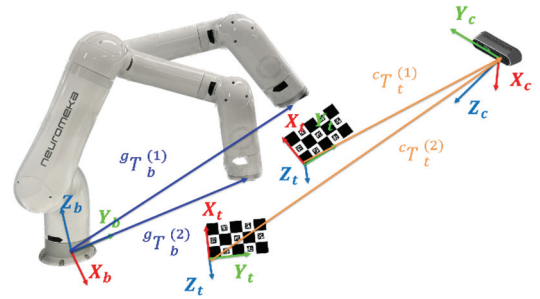


Fig. 5. First and Second Pose in Eye to Hand Calibration.

$$g_{T_b}^{(1)} b_{T_c} c_{T_t}^{(1)} = g_{T_b}^{(2)} b_{T_c} c_{T_t}^{(2)} \quad (3)$$

식(3)을 정리하게 되면 식 (4)와 같이 A와B 그리고 식 (5)와 같이 X로 표현될 수 있다.

$$\left(g_{T_b}^{(2)} \right)^{-1} g_{T_b}^{(1)} = A \quad (4)$$

$$c_{T_t}^{(2)} \left(c_{T_t}^{(1)} \right)^{-1} = B$$

$$b_{T_c} = X \quad (5)$$

Fig 6과 같이 다수의 위치에 해당하는 포즈에서 그리퍼와 타겟의 정보를 수집하고 저장한다. 또한 2D 이미지에서 3D로의 좌표 변환 시 동차변환행렬식(1) 성분 중 Z축인 t_{31} 값에 카메라의 Depth 정보를 통해 캘리브레이션 정확도를 검증한다.

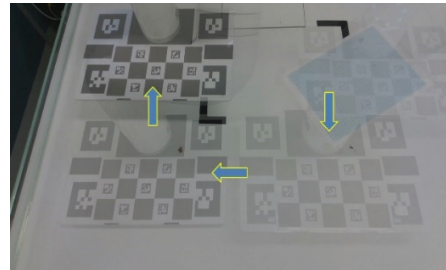


Fig. 6. Collecting Information of Gripper and Target Object Corresponding to Multiple Poses from 2D Images.

3. 실험 및 결과

3.1 실험 환경

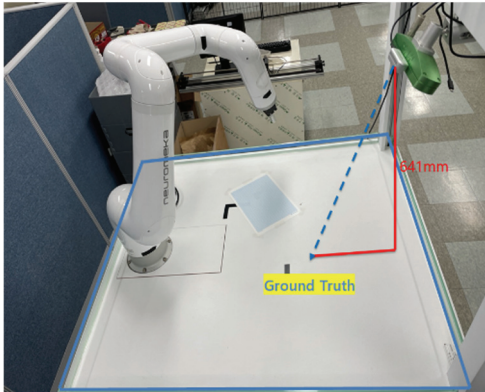
Table 1과 같은 사양을 가진 적외선 스테레오 비전 카메라와 Table 2와 같은 사양을 가진 로봇을 사용하여 실험을 진행하였다[15].

Table 1. Camera System Specification

Camera Model	Intel Realsense D435
Depth Sensor	Active IR Stereo
Depth Resolution	1280x720(Max)
Camera Resolution	1920x1080 30 fps
3D Camera Accuracy	<2% at 2m ²
Depth FOV	87° x 58°
RGB Sensor FOV	69° x 42°
Ideal Range	0.3m to 3m

Table 2. Camera System Specification

Model	Neurumecha Indy7
DOF	6 (all revolute)
Payload	7kg
Joint Motion Range	1,2,3,4,5:±175deg 6:±215deg
Maximum Joint Speed	1,2,3:150 4,5,6:180 deg/s
Maximum Tool Speed	1m/s
Maximum Reach	1.3m
Maximum Workspace w/Full Payload	0.8m
Repeatability	100μm

**Fig. 7.** Robot and Camera System.

실험 환경은 Fig 7 과 같이 로봇 원점과 같은 높이인 평면을 카메라가 촬영한다. 파란색으로 표시된 평면은 Ground Truth이고 파란색 점선은 카메라의 Z축이다. Ground Truth에서 카메라 까지의 높이는 641mm이다.

실험은 이미지 중심에 Charuco 보드를 위치시킨 후 X, Y

축으로 실제 100mm 증가하는 방향으로 이동시키고, Z축은 카메라와 가까워지는 방향(감소)으로 얼마나 정확히 측정되는지 실험한다.

3.2 카메라 캘리브레이션 정확성 실험

카메라 캘리브레이션 정확성은 2D 이미지를 통한 3D 좌표계 변환을 얼마나 정확히 했는지 실험한다. 2D 이미지에서 추정된 값과 2D + 3D(Depth) 값에서 추정된 값을 실제 측정된 값과 비교하는 실험을 진행하였다.

Table 3. After Camera Calibration, Measure Ground Truth and 2D Image

Axis	(Reference) Ground Truth	2D Image		
		Starting point	Ending point	Error(%) (Difference)
X-axis (mm)	0	4.98	108.94	3.96% (-103.96)
Y-axis (mm)	0	8.63	112.12	3.49% (-103.49)
Z-axis (mm)	641.00	532.13	462.45	30.32% (-69.68)
Roll(deg)	-22.16	-24.59	22.05	1.65% (-46.64)
Pitch(deg)	7.28	6.59	51.05	0.53% (-44.46)
Yaw(deg)	-2.53	-1.90	42.39	0.70% (-44.29)

Table 4. After Camera Calibration, Measure Ground Truth and 2D+3D(Depth) Image

Axis	(Reference) Ground Truth	2D + 3D(Depth)		
		Starting point	Ending point	Error(%) (Difference)
X-axis (mm)	0	0.42	104.12	3.70% (-103.7)
Y-axis (mm)	0	0.26	103.54	3.28% (-103.28)
Z-axis (mm)	641.00	649.00	545.00	4.00% (-4.00)
Roll(deg)	-22.16	-24.59	22.05	1.65% (-46.64)
Pitch(deg)	7.28	6.59	51.05	0.53% (-44.46)
Yaw(deg)	-2.53	-1.90	42.39	0.70% (-44.29)

실험 결과 2D 이미지의 경우 Table 3, Table 4과 같이 Ground Truth에 놓고 진행한 결과 실제 100mm 이동했을 때의 결과 중에서 X축과 Y축 이동은 큰 차이는 없었으나 Z축의 경우 시작점 높이부터 이동 후 높이까지 100mm를 이동해야 했으나 69.68mm를 이동하였고 2D + 3D(Depth)의 경우 Table 2 과 같이 시작점부터 이동 후 높이까지 오차는 4mm로 104mm를 이동하였다. 2D 이미지만으로 측정된 결과와 2D + 3D(Depth)로 측정된 결과 2D 이미지는 약 30% 오차를 보였고 2D + 3D(Depth)의 경우 약 4%의 오차를 보였다. 결론적으로 2D + 3D(Depth)가 보다 정확한 결과를 보여 주었다.

3.3 Eye to Hand 캘리브레이션 정확성 실험

Eye to hand 캘리브레이션은 카메라 캘리브레이션을 통해 얻은 타겟의 자세와 그리퍼의 자세를 이용한다. 카메라 캘리브레이션과 마찬가지로 이동정보는 매우 중요하므로 이동 및 회전 정보에 대해 2D 이미지에서 추정된 값과 2D + 3D(Depth) 값에서 추정된 값을 실제 측정된 값과 비교하는 실험을 진행한다.

Table 5. After Eye to Hand Calibration, Measure Ground Truth and 2D Image

Axis	(Reference) Ground Truth	2D Image	
		Result	Error(%) (Difference)
X-axis (mm)	462.90	436.63	5.67% (-26.27)
Y-axis (mm)	747.60	698.82	6.52% (-48.78)
Z-axis (mm)	641.00	536.53	16.29% (-104.47)
Roll(deg)	-22.16	-22.24	0.36% (-0.08)
Pitch(deg)	7.28	7.77	6.73% (-0.49)
Yaw(deg)	-2.53	-2.55	0.79% (-0.02)

실험 결과 2D 이미지만으로 Eye to Hand 캘리브레이션을 진행한 결과는 Table 5, Table 6과 같이 Ground Truth를 기준으로 Z축으로 104.47mm, X축과 Y축은 각각 26.27mm와 48.78mm의 차이가 났다. 이동 정보 중 Z값에 Depth 정보를 넣은 결과에서는 X축 25.59mm, Y축 50.22mm, Z축 6.99mm의 차이가 났다. 이 결과는 Z축 측정에서 2D이미지는 16.29% 오차가 발생했고 Depth 정보가 포함된 결과는 1.09%의 오차를 보였다. 이 결과를 보면 Z축 측정에서 매우 큰 차이를 보였고 로봇과 카메라의 캘리브레이션을 진행할

때 2D와 3D(Depth)의 정보를 결합하여 사용해야 더 정확히 물체 인식 및 파지가 가능할 것으로 판단된다.

Table 6. After Eye to Hand Calibration, Measure Ground Truth and 2D+3D(Depth) Image

Axis	(Reference) Ground Truth	2D + 3D(Depth)	
		Result	Error(%) (Difference)
X-axis (mm)	462.9	437.04	5.58% (-25.59)
Y-axis (mm)	747.6	697.38	6.71% (-50.22)
Z-axis (mm)	641	634.01	1.09% (-6.99)
Roll(deg)	-22.16	-22.32	0.72% (-0.16)
Pitch(deg)	7.28	7.86	7.96% (-0.58)
Yaw(deg)	-2.53	-2.56	1.18% (-0.03)

4. 결 론

본 연구에서는 반도체 자동화를 위한 빈피킹 로봇의 비전 기반 캘리브레이션 방법에 관한 연구를 진행하였다. 로봇과 타겟, 카메라 좌표계의 회전과 이동 변환에 대해 정의하였다.

2D 이미지를 통한 3D 좌표 복원은 카메라 캘리브레이션이 진행되어야 정확한 이미지를 얻을 수 있다. 3D 좌표 변환은 거리 값을 추정하기 어려우므로 RGB-D 카메라의 Depth 정보를 이동 정보의 Z값에 사용하였다. 2D 이미지를 통해 변환한 결과와 2D + 3D(Depth)정보를 이용한 결과를 Ground Truth와 비교하여 정확성 평가 실험을 진행한 결과 Z축 측정에서 2D 이미지는 약 30% 오차를 보였고 2D+3D(Depth)의 경우 약 4%의 오차를 보였다.

Eye to Hand 캘리브레이션에서도 2D 이미지에서 추정한 거리 값과 2D + 3D(Depth)에서 추정한 값을 Ground Truth와 비교하여 정확성 평가 실험을 진행하였다. 그 결과 Z축 측정에서 2D이미지만 사용한 경우 16.29% 오차가 발생했고 Depth 정보가 포함된 결과는 1.09%의 오차를 보여 Eye to Hand 캘리브레이션 정확도를 높이기 위해서는 Depth 정보의 활용이 중요한 것을 확인하였다. Depth 카메라를 사용함으로써 더 정확한 물체 인식 및 파지가 가능할 것으로 판단된다.

향후 연구로는 2D 비전과 Point 클라우드를 이용하여 캘리브레이션을 진행해 볼 필요가 있다. Point 클라우드는 현실을 더 정확히 파악하여 3D로 변환하기 때문에 좀 더 정밀한 측정이 가능할 것으로 예상된다.

감사의 글

이 논문은 경기도의 경기도협력연구센터(GRRC)사업 [(GRRC TU Korea2020-B02), 이종소재 접합 제조공정 자동화를 위한 로봇 응용기술 개발]과 2022년도 정부(산업통상자원부와 한국산업기술진흥원의 '한/체코 국제공동 기술개발사업(No. P0019623)'으로 수행된 연구 결과입니다.

참고문헌

1. Torres, Pedro, et al. "Bin-Picking Solution for Randomly Placed Automotive Connectors Based on Machine Learning Techniques." *Electronics* 11.3 (2022): 476.
2. Lee, Sukhan, and Yeonho Lee. "Real-Time Industrial Bin-Picking with a Hybrid Deep Learning-Engineering Approach." 2020 IEEE International Conference on Big Data and Smart Computing (BigComp). IEEE, 2020.
3. Ellekilde, Lars-Peter, and Henrik Gordon Petersen. "Motion planning efficient trajectories for industrial bin-picking." *The International Journal of Robotics Research* 32.9-10 (2013): 991-1004.
4. Rahardja, Krisnawan, and Akio Kosaka. "Vision-based bin-picking: Recognition and localization of multiple complex objects using simple visual cues." *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems. IROS'96. Vol. 3. IEEE, 1996.*
5. Ha, Eun-Hyeon, and Goon-Ho Choi. "Position Control of an Object Using Vision Sensor." *Journal of the Semiconductor & Display Technology* 10.2 (2011): 49-56.
6. Lee, Yong-Hwan, and Youngseop Kim. "Implementation of object feature extraction within image for object tracking." *Journal of the Semiconductor & Display Technology* 17.3 (2018): 113-116.
7. Park, Frank C., and Bryan J. Martin. "Robot sensor calibration: solving $AX=XB$ on the Euclidean group." *IEEE Transactions on Robotics and Automation* 10.5 (1994): 717-721.
8. Open-Source Computer Vision 4.6.0, "Camera Calibration and 3D Reconstruction", Generated on Sun Jun 5, 2022, 16:19:58 for OpenCV, 2022.
9. Fassi, Irene, and Giovanni Legnani. "Hand to sensor calibration: A geometrical interpretation of the matrix equation $AX=XB$." *Journal of Robotic Systems* 22.9 (2005): 497-506.
10. Li, Haiyuan, et al. "Simultaneous hand-eye and robot-world calibration by solving the $AX= YB$ problem without correspondence." *IEEE Robotics and Automation Letters* 1.1 (2015): 145-152.
11. Condurache, D., and A. Burlacu. "Orthogonal dual tensor method for solving the $AX=XB$ sensor calibration problem." *Mechanism and Machine Theory* 104 (2016): 382-404.
12. Horaud, Radu, and Fadi Dornaika. "Hand-eye calibration." *The international journal of robotics research* 14.3 (1995): 195-210.
13. Tabb, Amy, and Khalil M. Ahmad Yousef. "Solving the robot-world hand-eye (s) calibration problem with iterative methods." *Machine Vision and Applications* 28.5 (2017): 569-590.
14. Park, Frank C., and Bryan J. Martin. "Robot sensor calibration: solving $AX=XB$ on the Euclidean group." *IEEE Transactions on Robotics and Automation* 10.5 (1994): 717-721.
15. Lee, Yong-Hwan, and Youngseop Kim. "Comparative Analysis of Cost Aggregation Algorithms in Stereo Vision." *Journal of the Semiconductor & Display Technology* 15.1 (2016): 47-51.

접수일: 2023년 3월 3일, 심사일: 2023년 3월 16일,
게재확정일: 2023년 3월 22일