

Comparative Analysis of Depth Sampling Strategies on Coordinate Transformation Consistency in a ROS-Based RGB-D Pipeline

Kwang-Ho Yang*, Hyeon-O Choe**

*Researcher, Major in Smart Agriculture, Suncheon National University, Suncheon, Korea

**Professor of Academic Research, Low-Carbon Agriculture-Based Smart Distribution Research Center, Suncheon National University, Suncheon, Korea

[Abstract]

This paper comparatively analyzes depth sampling strategies and their effect on coordinate transformation consistency in a ROS-based RGB-D perception pipeline. The proposed pipeline consists of RGB-based ROI detection, window-based depth sampling for 3D position estimation, ROS TF coordinate transformation, and TCP-based control command delivery. Three sampling strategies (1×1 single-pixel, 3×3 and 5×5 window median) were evaluated across nine experimental conditions, including three camera-to-object distances (30, 40, 50 cm) and three lighting levels (bright, normal, dim), totaling 270 repeated trials. The 3×3 window median improved consistency by 5.7 times over single-pixel sampling, and 5×5 by 8.1 times (Kruskal-Wallis $H=71.59$, $p<0.001$). All conditions maintained deviations below 1.3 mm, confirming that depth sampling strategy is a critical design parameter for coordinate transformation reliability.

▶ **Key words:** RGB-D Camera, Coordinate Transformation, ROS, Depth Sampling, Consistency Evaluation

[요 약]

본 논문에서는 ROS 기반 RGB-D 인식 파이프라인에서 깊이 샘플링 전략이 좌표 변환 일관성에 미치는 영향을 비교 분석하였다. 제안한 파이프라인은 RGB 영상 기반 관심영역 검출, 윈도우 기반 깊이 샘플링을 통한 3차원 위치 산출, ROS TF 좌표 변환, TCP 제어 명령 전달로 구성된다. 1×1 단일 픽셀, 3×3, 5×5 윈도우 중앙값의 세 가지 샘플링 전략을 카메라-대상 거리 3단계(30, 40, 50 cm)와 조명 조건 3단계(밝음, 보통, 어두움)를 포함한 9개 조건에서 총 270회 반복 실험하였다. 3×3 방식은 단일 픽셀 대비 5.7배, 5×5는 8.1배 일관성이 개선되었으며(Kruskal-Wallis $H=71.59$, $p<0.001$), 모든 조건에서 1.3 mm 이하의 변동을 유지하였다. 이를 통해 깊이 샘플링 전략이 좌표 변환 일관성의 핵심 설계 파라미터임을 확인하였다.

▶ **주제어:** RGB-D 카메라, 좌표 변환, ROS, 깊이 샘플링, 일관성 분석

-
- First Author: Kwang-Ho Yang, Corresponding Author: Hyeon-O Choe
 - *Kwang-Ho Yang (yg8686@naver.com), Major in Smart Agriculture, Suncheon National University
 - **Hyeon-O Choe (wishind@scnu.ac.kr), Low-Carbon Agriculture-Based Smart Distribution Research Center, Suncheon National University
 - Received: 2026. 01. 26, Revised: 2026. 04. 13, Accepted: 2026. 04. 15.

I. Introduction

최근 로봇 시스템은 제조, 물류, 서비스 등 다양한 분야에서 센서 기반 인식 결과를 실시간으로 제어 명령으로 연계하는 방향으로 발전하고 있다[1-2]. 이러한 시스템에서는 카메라나 거리 센서를 통해 획득한 인식 결과를 로봇 제어 좌표계로 변환하는 과정이 필수이며, 변환 과정의 신뢰성은 반복 동작의 안정성과 직접적으로 연결된다. 특히 반복 실행 환경에서는 작은 좌표 변환 변동이 누적되어 제어 품질을 저하시킬 수 있으므로, 좌표 변환 결과의 일관성 평가가 중요하다.

RGB-D 카메라는 색상 정보(RGB)와 깊이 정보(Depth)를 동시에 제공함으로써 3차원 위치 산출에 널리 활용되어 왔다[3-4]. 그러나 RGB-D 센서의 깊이 측정은 거리에 따라 노이즈가 이차적으로 증가하고, 조명 조건이나 입사각에 의해서도 영향을 받는 것으로 보고되고 있다[5]. 특히 Intel RealSense D435 계열 카메라에 대한 노이즈 분석 연구에서는 거리 및 각도에 따른 축방향-횡방향 노이즈 특성이 실증적으로 모델링된 바 있으며[5], 로봇 응용에서의 깊이 정확도 비교 연구에서도 1m 이내 근거리에서 1cm 이하의 오차가 보고되었다[6]. 이러한 깊이 노이즈는 단일 픽셀 기반 좌표 산출 시 프레임마다 미세한 변동을 유발할 수 있으며, 이를 완화하기 위한 필터링 기법으로는 중앙값 필터, 가우시안 필터 등이 포인트 클라우드 처리 분야에서 연구되어 왔다[7].

로봇 소프트웨어 측면에서 ROS는 센서 데이터 처리, 제어, 통신을 모듈화하여 통합할 수 있는 표준 미들웨어로 널리 활용되고 있으며[8], 좌표계 변환을 위해 TF/TF2 라이브러리를 제공한다[9]. 기존 연구에서는 ROS 기반 좌표 변환을 활용하여 인식 결과를 로봇 제어에 연계하는 시스템이 다수 제안되었으나[10-12], 이러한 연구들은 주로 작업 성공률이나 인식 정확도 중심으로 시스템 성능을 평가하였으며, 좌표 변환 과정에서 발생하는 변동이나 반복 실행 시의 일관성에 대한 정량적 평가는 제한적으로 이루어져 왔다.

이에 본 논문에서는 ROS 기반 RGB-D 인식 좌표 변환 파이프라인에서 깊이 샘플링 전략이 좌표 변환 일관성에 미치는 영향을 비교 분석한다. 구체적으로, 단일 픽셀(1×1), 3×3 윈도우 중앙값, 5×5 윈도우 중앙값의 세 가지 깊이 샘플링 전략을 비교하고, 카메라-대상 거리와 조명 조건을 변화시킨 총 9개 실험 조건에서 270회의 반복 실험을 수행하여 좌표 변환 결과의 일관성을 정량적으로 분석한다. 본 연구는 객체 검출 정확도나 작업 성공률을 직접 평가하

기보다, 인식-제어 연계 과정에서 좌표 변환의 설계 파라미터와 환경 요인이 일관성에 미치는 영향을 독립적으로 검증한다는 점에서 기존 연구와 차별성을 가진다.

II. Related works

1. RGB-D-Based 3D Position Estimation and Perception

RGB-D 카메라는 RGB 영상과 깊이 정보를 동시에 제공함으로써 객체의 3차원 위치를 직접적으로 추정할 수 있다는 장점을 가진다. 이러한 특성으로 인해 RGB-D 센서는 객체 인식, 위치 추정, 3차원 환경 인식 등 다양한 로봇 응용 분야에서 활용되어 왔다[3-4]. 특히 깊이 정보를 활용한 3차원 좌표 산출 방식은 단일 RGB 영상 기반 접근법에 비해 거리 정보의 불확실성을 줄일 수 있어, 로봇 제어 연계에 적합한 입력 정보를 제공한다.

그러나 RGB-D 센서의 깊이 측정은 거리, 입사각, 조명 조건 등에 의해 노이즈가 발생하며, Ahn 등[5]은 Intel RealSense D435의 축방향 및 횡방향 노이즈가 거리의 제곱에 비례하여 증가함을 실증적으로 모델링하였다. Rustler 등[6]은 D435를 포함한 4종의 스테레오 깊이 카메라를 비교하여, 1m 이내 근거리에서 1cm 이하의 오차를 보고하였다. 이러한 연구들은 RGB-D 센서의 깊이 측정 특성을 규명하였으나, 깊이 노이즈가 좌표 변환 파이프라인의 반복 일관성에 미치는 영향을 분석한 사례는 제한적이다.

2. Depth Sampling and Point Cloud Filtering

깊이 센서의 노이즈를 완화하기 위한 필터링 기법은 포인트 클라우드 처리 분야에서 다양하게 연구되어 왔다[7]. 대표적으로 가우시안 필터는 주변 포인트의 가중 평균을 통해 랜덤 노이즈를 저감하는 데 효과적이나, 경계 영역에서 평활화가 발생할 수 있다. 반면 중앙값 필터는 이상치에 대한 강건성이 높아 깊이 센서의 간헐적 오측정에 적합한 것으로 알려져 있다[7].

그러나 이러한 필터링 연구들은 주로 포인트 클라우드 전체의 품질 개선에 초점을 맞추고 있으며, 특정 관심 좌표에 대한 깊이 샘플링 윈도우 크기가 좌표 변환 일관성에 미치는 영향을 정량적으로 비교한 연구는 부족한 실정이다.

3. ROS-Based Coordinate Transformation and Control Integration

ROS는 센서 데이터 처리, 로봇 제어, 통신을 모듈화하여 통합할 수 있는 로봇 소프트웨어 플랫폼으로 널리 활용되고 있다[8]. ROS 환경에서 좌표 변환은 TF 및 TF2 라이브러리를 통해 관리되며, TF2는 다수의 좌표 프레임 간 변환 관계를 시간 정보와 함께 관리하고, 특정 시점의 변환을 조회하여 점이나 벡터를 다른 좌표계로 변환할 수 있도록 설계되었다[9].

기존 연구에서는 ROS 기반 좌표 변환을 활용하여 인식 결과를 로봇 제어에 연계하는 다양한 시스템 구조가 제안되었다[10-12]. 그러나 이러한 연구들에서는 주로 전체 시스템의 동작 여부나 작업 성공률을 중심으로 결과를 제시하는 경우가 많으며, 좌표 변환 과정에서 발생하는 변동이나 반복 실행 시의 일관성에 대한 정량적 평가는 제한적으로 이루어져 왔다. 특히 좌표 변환 정확도를 절대 기준과 비교하기보다는, 동일 조건 반복 실험에서의 내부 변동 특성을 분석한 연구는 충분하지 않다.

4. Summary and Differentiation of This Study

기존 연구들은 크게 세 가지 관점에서 한계를 보인다. 첫째, RGB-D 센서의 깊이 노이즈 연구[5-6]는 센서 자체의 측정 특성 규명에 집중되어 있으며, 노이즈가 로봇 제어 좌표 변환의 반복 일관성에 미치는 영향은 다루지 않았다. 둘째, 포인트 클라우드 필터링 연구[7]는 전체 데이터의 품질 개선을 목표로 하며, 관심 좌표에 대한 샘플링 윈도우 크기 비교는 수행되지 않았다. 셋째, ROS 기반 로봇 응용 연구[10-12]는 작업 성공률 중심의 평가에 치중하여, 좌표 변환 단계 자체의 일관성을 분리 검증하지 않았다.

본 연구는 이러한 공백을 보완하기 위해, 깊이 샘플링 전략(1×1, 3×3, 5×5)을 비교하고 카메라-대상 거리 및 조명 조건 변화를 포함한 다조건 반복 실험을 통해 좌표 변환 파이프라인의 일관성을 정량적으로 분석한다는 점에서 차별성을 가진다.

III. Proposed Coordinate Transformation Pipeline

본 장에서는 RGB-D 기반 인식 결과를 로봇 제어 좌표로 변환하기 위해 본 연구에서 구성한 좌표 변환 파이프라인의 구조와 처리 절차를 설명한다. 제안한 파이프라인은 RGB 영상 기반 인식 결과와 깊이 정보를 결합하여 3차원

좌표를 산출하고, 이를 ROS 환경에서 로봇 기준 좌표계로 변환한 후 제어 입력으로 전달하는 과정으로 구성된다. 전체 파이프라인의 개요는 Fig. 4에 나타내었다.

1. System Overview

본 연구에서 사용한 시스템은 RGB-D 카메라, 로봇 플랫폼, 그리고 ROS 기반 처리 노드들로 구성된다. RGB-D 카메라는 로봇 엔드이펙터 상부에 고정되어 로봇 작업 공간을 관측하며, 본 연구에서는 Intel RealSense D435i 카메라를 사용하였다[13]. 로봇에 장착된 카메라 모듈의 실제 구성은 Fig. 1에 나타나 있다.



Fig. 1. RGB-D camera module(D435i) mounted on the robot

실험 환경은 연구용 수직농장 모형으로 구성되었으며, 인식 대상은 수직 베드에 고정된 딸기 모형이다. 수직농장 베드에 설치된 딸기 모형의 배치 예시는 Fig. 2에 제시하였으며, 전체 실험 환경과 로봇 시스템의 배치는 Fig. 3에 나타내었다.



Fig. 2. Strawberry mock-up installed on the vertical-farm bed



Fig. 3. Experimental setup consisting of the research vertical-farm mock-up and the robot system

전체 시스템은 ROS 환경에서 구동되며, RGB 영상 수신, 인식 결과 처리, 깊이 기반 좌표 산출, 좌표 변환, 그리고 로봇 제어 명령 전달이 각각 독립적인 노드로 구성되어 토픽 및 서비스 기반으로 연동된다. 카메라 프레임과 로봇 기준 프레임 간의 정적 변환 관계는 ROS의 `static_transform_publisher`를 이용하여 설정하며, TF lookup은 `rospy.Time(0)`을 사용하여 최신 가용 변환을 조회한다.

2. 3D Coordinate Computation from RGB-D Data

RGB 영상에서 객체가 인식되면, 인식 노드는 객체 중심에 해당하는 2차원 픽셀 좌표 (u, v) 를 산출한다. 해당 좌표는 RGB 영상과 정합된 깊이 영상 상의 동일 위치에 대응되며, 깊이 영상으로부터 깊이 값 $Z(u, v)$ 를 획득한다.

카메라 좌표계 기준의 3차원 좌표 (X_c, Y_c, Z_c) 는 카메라 내부 파라미터 (f_x, f_y, c_x, c_y) 를 이용하여 계산된다. 이때 3차원 좌표 산출식은 다음과 같다.

$$X_c = \frac{(u - c_x) \cdot Z(u, v)}{f_x} \quad (1)$$

$$Y_c = \frac{(v - c_y) \cdot Z(u, v)}{f_y} \quad (2)$$

$$Z_c = Z(u, v) \quad (3)$$

식 (1)-(3)은 RGB-D 카메라 모델에 기반한 기본적인 투영 관계를 이용하여 픽셀 좌표를 카메라 좌표계의 3차원 위치로 변환하는 과정이다. 이 과정에서 사용되는 카메라 내부

파라미터는 사전 보정된 값을 사용하였으며, RGB 영상과 깊이 영상은 동일한 좌표계로 정합된 상태에서 처리된다.

3. Window-based Depth Sampling

기본적인 깊이 기반 좌표 산출에서는 인식 결과로 도출된 단일 픽셀 좌표 (u, v) 에 대응하는 하나의 깊이 값을 사용한다. 그러나 앞서 언급한 바와 같이 RGB-D 센서의 깊이 측정은 프레임마다 미세한 변동을 보일 수 있으며[5], 이는 좌표 변환 결과의 일관성에 직접적인 영향을 미친다.

본 연구에서는 깊이 샘플링 전략이 좌표 변환 일관성에 미치는 영향을 비교하기 위해, 단일 픽셀(1×1) 방식 외에 윈도우 기반 샘플링 방식을 도입하였다. 윈도우 기반 방식에서는 인식 결과 좌표 (u, v) 를 중심으로 $N \times N$ 크기의 윈도우 내 유효 깊이 값들을 추출한 후, 중앙값(median)을 대표 깊이 값으로 사용한다. 중앙값은 평균값 대비 이상치에 대한 강건성이 높아[7], 깊이 센서에서 간헐적으로 발생하는 비정상적인 측정값의 영향을 제한할 수 있다.

본 연구에서는 1×1(단일 픽셀, 기존 방식), 3×3, 5×5의 세 가지 윈도우 크기를 비교 대상으로 설정하였으며, 각 윈도우에서의 유효 깊이 값 수와 윈도우 내 깊이 표준편차를 함께 기록하여 샘플링 전략에 따른 차이를 정량적으로 분석하였다.

4. Coordinate Transformation Using ROS TF

카메라 좌표계 기준으로 계산된 3차원 위치는 로봇 제어에 직접 활용할 수 없으므로, 로봇 기준 좌표계로의 변환이 필요하다. 본 연구에서는 ROS에서 제공하는 TF 프레임워크를 이용하여 좌표 변환을 수행하였다[9]. 카메라 프레임과 로봇 기준 프레임 간의 관계는 회전 행렬과 병진 벡터로 구성된 변환 행렬로 표현된다.

카메라 좌표계 기준 점 $P_c = [X_c, Y_c, Z_c, 1]^T$ 는 다음과 같이 로봇 기준 좌표계의 점 P_b 로 변환된다.

$$P_b = T_{camera}^{base} P_c = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \quad (4)$$

본 연구에서는 ROS TF2 라이브러리를 이용하여 해당 변환 행렬을 실시간으로 조회하고, 인식 결과에 대응되는 3차원 좌표를 로봇 기준 좌표계로 변환하였다. TF lookup은 `rospy.Time(0)`을 사용하여 최신 가용 변환을 조회하며, 실험 결과 TF lookup 소요시간은 평균 0.1~0.3ms 수

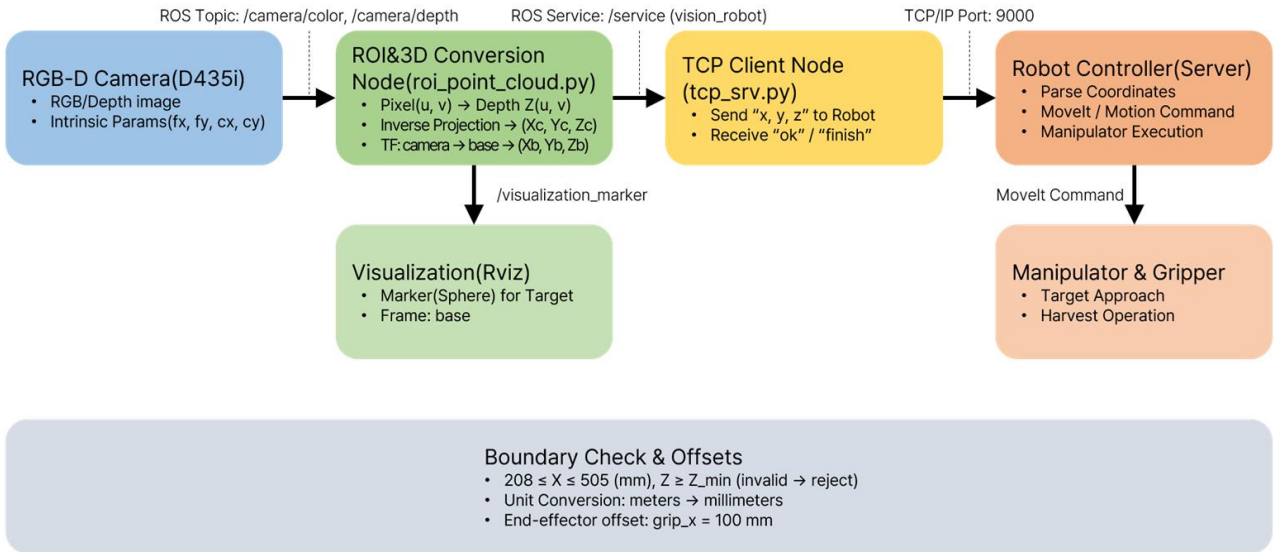


Fig. 4. Pipeline structure for coordinate transformation and robot-control integration using RGB-D perception results

준으로 확인되었다. 이러한 좌표 변환 및 제어 연계 구조의 전체 흐름은 Fig. 4에 나타나 있다.

5. Control Coordinate Normalization

로봇 기준 좌표계로 변환된 위치 정보는 로봇 제어 시스템에서 요구하는 입력 형식에 맞게 정규화된다. 본 연구에서는 좌표 단위를 미터 단위에서 밀리미터 단위로 변환하고, 제어 기준 프레임과의 정렬을 위해 일정한 기준 오프셋을 적용하였다.

$$X = 1000 \cdot X_b - \Delta_x \quad (5)$$

$$Y = 1000 \cdot Y_b \quad (6)$$

$$Z = 1000 \cdot Z_b \quad (7)$$

이 정규화 과정은 실제 작업 수행을 위한 보정이 아니라, 반복 실험에서 좌표 변환 결과를 일관되게 해석하기 위한 기준 설정을 목적으로 한다.

6. Failure Classification System

좌표 변환 파이프라인의 실패 사례를 체계적으로 분석하기 위해, 본 연구에서는 파이프라인 실행 과정에서 발생할 수 있는 실패를 다음 5개 카테고리 분류하였다. (1) 미검출(No Detection): 인식 노드에서 대상 객체를 검출하지 못한 경우. (2) 깊이 NaN(Depth NaN): ROI 좌표에 대응하는 깊이 값이 유효하지 않은 경우로, 깊이 센서의 측정 범위 초과나 반사면에 의한 깊이 계산 실패 등이 원

인이 될 수 있다. (3) TF Lookup 실패: ROS TF를 통한 좌표 변환 조희가 실패한 경우로, 노드 초기화 시 시간 동기화 지연이 주요 원인이다. (4) 범위 초과(Boundary Rejection): 좌표 변환은 성공하였으나, 변환된 제어 좌표가 로봇의 유효 작업 범위를 초과한 경우. (5) 서비스 오류(Service Failure): TCP 통신을 통한 로봇 제어 서비스 호출이 실패한 경우이다. 이 분류 체계를 통해 실패 원인을 파이프라인의 특정 단계에 대응시킬 수 있으며, 각 실험 조건에서의 실패 분포를 정량적으로 분석할 수 있다.

IV. Experiments and Results

1. Experimental Setup and Conditions

본 연구에서는 제안한 좌표 변환 파이프라인의 특성을 다각적으로 분석하기 위해 3가지 실험 세트를 설계하였다. 모든 실험에서 동일한 하드웨어(Intel RealSense D435i, 6축 로봇 매니플레이터)와 소프트웨어(ROS Noetic, TF2) 구성을 사용하였다. 또한, 일관성 실험의 특성상, 매 시행마다 동일한 대상의 동일한 좌표가 반복 측정되어야 한다. 이를 위해 인식 대상을 단일 딸기 모형 1개로 한정하여, 인식 노드가 매 프레임 동일한 객체의 중심 좌표를 출력하도록 구성하였다. 로봇 팔은 인식된 좌표를 기반으로 접근 동작을 수행하였으나, 수확 등의 실제 작업은 수행하지 않았다.

실험 1(깊이 샘플링 비교)에서는 1×1, 3×3, 5×5의 세 가지 깊이 샘플링 윈도우 크기에 대해 각 30회씩 총 90회의 반복 실험을 수행하였다. 카메라-대상 거리와 조명 조건은 동일하게 유지하였다.

실험 2(거리 변화)에서는 카메라와 인식 대상 간 거리를 30cm, 40cm, 50cm의 3단계로 변경하면서 각 30회씩 총 90회의 실험을 수행하였다. 깊이 샘플링은 3×3 윈도우 중앙값 방식을 사용하였다.

실험 3(조명 변화)에서는 밝음(전체 조명), 보통(간접 조명), 어두움(최소 조명)의 3가지 조명 조건에서 각 30회씩 총 90회의 실험을 수행하였다.

전체적으로 9개 조건에서 총 270회의 반복 실험을 수행하였으며, 각 시행에서 ROI 좌표, 카메라 좌표계 3차원 위치, 로봇 기준 좌표계 위치, 정규화된 제어 좌표, 깊이 윈도우 표준편차, TF lookup 소요시간을 CSV 형태로 기록하였다. 좌표 변환 결과의 일관성은 각 조건에서의 성공 사례에 대해 평균점으로부터의 유클리드 거리를 산출하여 평가하였으며, Shapiro-Wilk 검정을 통해 정규성을 검정하고, 비모수 검정(Kruskal-Wallis, Mann-Whitney U)을 통해 조건 간 차이의 통계적 유의성을 평가하였다. 실험 조건의 요약은 Table 1에 제시하였다.

Table 1. Summary of experimental conditions

Experiment	Condition	Sampling	Trials	Success	Fail	Rate (%)
Exp.1 (Sampling)	1x1 (baseline)	1x1	30	30	0	100.0
	3x3	3x3	30	30	0	100.0
	5x5	5x5	30	29	1	96.7
Exp.2 (Distance)	30cm	3x3	30	29	1	96.7
	40cm	3x3	30	29	1	96.7
	50cm	3x3	30	30	0	100.0
Exp.3 (Lighting)	Bright	3x3	30	30	0	100.0
	Normal	3x3	30	30	0	100.0
	Dim	3x3	30	29	1	96.7
Total			270	266	4	98.5

2. Depth Sampling Comparison Results

동일한 인식 대상과 환경 조건에서 깊이 샘플링 윈도우 크기에 따른 좌표 변환 일관성 비교 결과를 Table 2에 나타내었다.

Table 2. Coordinate consistency by depth sampling window size

Sampling	X std (mm)	Y std (mm)	Z std (mm)	Euclid. mean (mm)	Euclid. std (mm)
1x1	0.499	0.238	0.252	0.580	0.142
3x3	0.183	0.077	0.133	0.102	0.211
5x5	0.186	0.078	0.004	0.072	0.184

Table 2에서 확인할 수 있듯이, 1×1 단일 픽셀 샘플링의 평균 유클리드 변동은 0.580mm인 반면, 3×3 윈도우 중앙값 방식은 0.102mm로 약 5.7배, 5×5는 0.072mm로

약 8.1배 개선되었다. 특히 Z축(깊이 방향) 표준편차에서 5×5 윈도우는 0.004mm로 1×1(0.252mm) 대비 현저한 안정화를 보였다. 이는 윈도우 크기가 커질수록 깊이 센서의 프레임 간 노이즈가 중앙값 연산에 의해 효과적으로 억제됨을 의미한다.

Kruskal-Wallis 검정 결과 세 그룹 간 유의한 차이가 확인되었으며(H=71.59, p<0.001), 사후 검정(Mann-Whitney U)에서도 1×1 vs 3×3(p<0.001), 1×1 vs 5×5(p<0.001), 3×3 vs 5×5(p<0.001) 모든 쌍별 비교에서 통계적으로 유의한 차이가 나타났다. 이 결과는 깊이 샘플링 전략이 좌표 변환 파이프라인의 일관성에 실질적인 영향을 미치는 핵심 설계 파라미터임을 시사한다. 샘플링 윈도우 크기별 평균 유클리드 변동의 비교는 Fig. 5에 나타내었다.

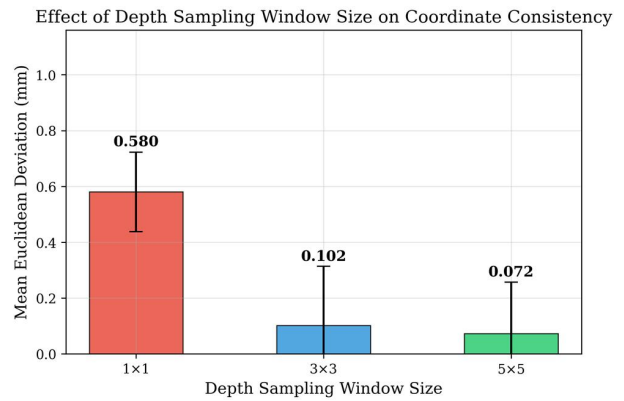


Fig. 5. Effect of depth sampling window size on coordinate consistency

3. Distance Variation Results

카메라-대상 거리 변화에 따른 좌표 변환 일관성 분석 결과를 Table 3에 정리하였다. 모든 거리 조건에서 3×3 윈도우 중앙값 샘플링을 적용하였다.

Table 3. Coordinate consistency by camera-to-object distance

Distance	X std (mm)	Y std (mm)	Z std (mm)	Euclid. mean (mm)	Depth std (mm)
30cm	0.310	0.130	0.099	0.217	0.406
40cm	0.713	0.309	0.369	0.794	0.544
50cm	0.718	0.220	0.030	0.626	0.314

30cm 거리에서 0.217mm, 40cm에서 0.794mm, 50cm에서 0.626mm의 평균 유클리드 변동을 보였다. Table 3의 Depth std는 각 시행에서 3×3 샘플링 윈도우 내 깊이 값들의 표준편차를 전체 성공 시행에 대해 평균한 값으로, 해당 위치에서의 깊이 측정 안정성을 나타낸다. 윈도우 내 인접 픽셀들은 동일 표면의 매우 좁은 영역(약 1~2mm²)을 관측

하므로 깊이 값의 차이가 작으며, 0.3~0.5mm 수준의 Depth std는 센서의 프레임 간 미세 노이즈에 해당한다. 전반적으로 거리가 증가할수록 변동이 커지는 경향을 보이며, 이는 RGB-D 센서의 깊이 오차가 거리의 제곱에 비례하여 증가한다는 기존 연구[5]와 일관된 결과이다.

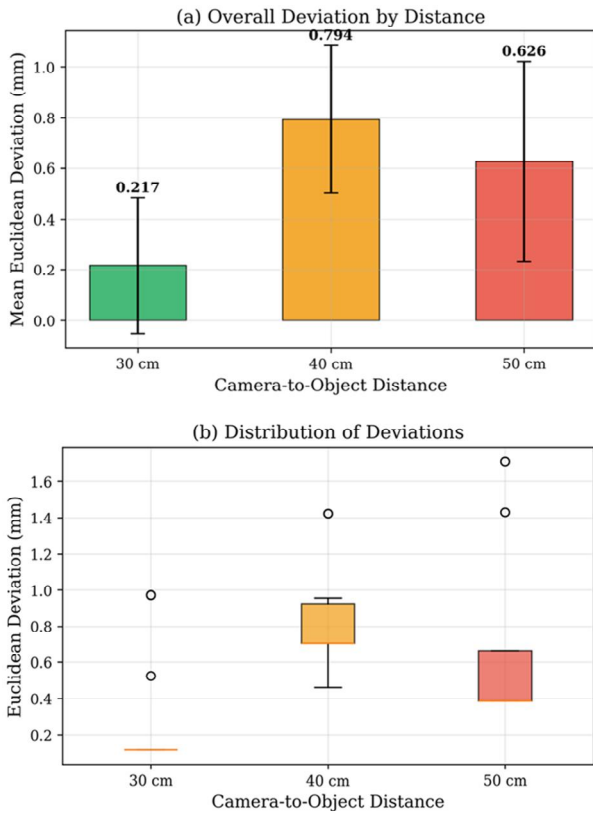


Fig. 6. Effect of camera-to-object distance on coordinate consistency: (a) overall deviation, (b) distribution

50cm에서 40cm보다 변동이 다소 작게 나타난 것은 Z 축 표준편차가 0.030mm로 현저히 낮아진 영향이다. 이는 원거리에서 깊이 해상도가 낮아지면서 양자화 단위가 커지고, 특정 거리에서 깊이 값이 동일한 이산 값에 수렴하는 양자화 특성에 기인하는 것으로 분석된다. 거리별 평균 변동 및 분포 특성은 Fig. 6에 나타내었다.

4. Lighting Variation Results

조명 조건 변화에 따른 좌표 변환 일관성 분석 결과를 Table 4에 정리하였다.

Table 4. Coordinate consistency by lighting condition

Lighting	X std (mm)	Y std (mm)	Z std (mm)	Euclid. mean (mm)	Depth std (mm)
Bright	1.281	0.429	0.600	1.248	0.854
Normal	0.450	0.139	0.375	0.486	0.511
Dim	0.258	0.080	0.417	0.441	0.000

밝은 조명 조건에서 1.248mm로 가장 큰 변동을 보인 반면, 보통 조명에서 0.486mm, 어두운 조명에서 0.441mm의 변동을 보였다. 이 결과는 직관적 예상과 달리 밝은 환경이 오히려 좌표 변환 일관성을 저하시키는 것을 보여준다.

이는 Intel RealSense D435i의 깊이 센서가 IR(적외선) 패턴 투사 방식을 사용하기 때문으로 분석된다. 강한 가시광 환경에서는 IR 패턴에 대한 간섭이 증가하여 깊이 측정의 변동성이 커지며, 이러한 분석은 Table 4의 Depth std에서도 뒷받침된다. 밝은 조건(0.854mm)이 보통(0.511mm) 및 어두운(0.000mm) 조건보다 높게 나타나며, 이는 밝은 환경에서 동일 윈도우 내 인접 픽셀 간에도 깊이 값의 변동이 증가함을 의미한다. 어두운 조건에서 Depth std가 0.000mm로 나타난 것은 IR 프로젝터의 패턴이 외부 간섭 없이 안정적으로 투사되어 윈도우 내 깊이 값이 균일하게 측정되었기 때문으로 판단된다. 이러한 결과는 RGB-D 기반 파이프라인을 설계할 때 조명 조건이 깊이 측정 안정성에 미치는 영향을 고려해야 함을 시사한다. 조명 조건별 변동 비교 및 분포 특성은 Fig. 7에 나타내었다.

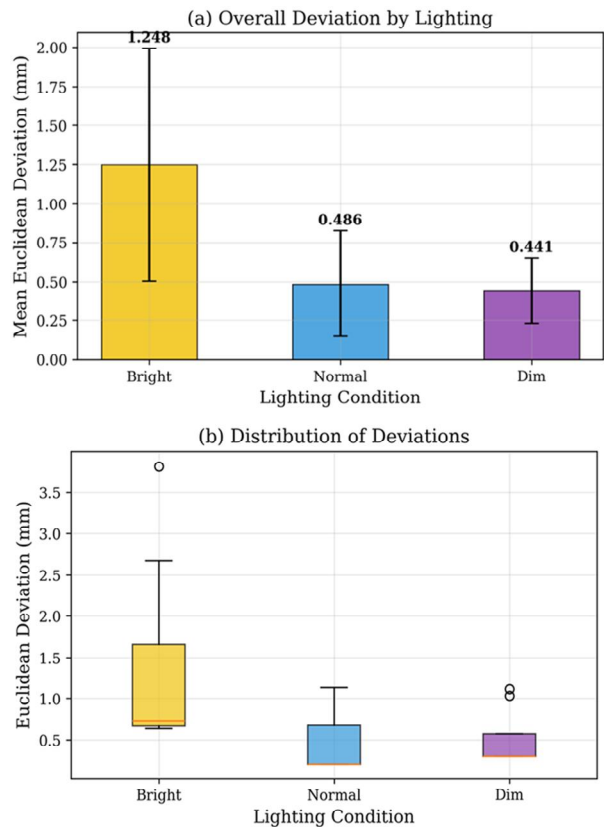


Fig. 7. Effect of lighting conditions on coordinate consistency: (a) overall deviation, (b) distribution

5. Failure Analysis

전체 270회 실험에서 266회(98.5%)가 성공하였으며, 4회(1.5%)가 실패하였다. 실패 사례는 모두 TF lookup 실패(fail_tf_lookup)로 분류되었으며, 미검출, 깊이 NaN, 범위 초과, 서비스 오류에 해당하는 실패는 관찰되지 않았다.

TF lookup 실패는 모든 경우 각 실험 조건의 노드 초기화 직후 첫 번째 시행에서 발생하였으며, 이는 ROS TF 라이브러리의 변환 정보 버퍼가 충분한 데이터를 축적하기 전에 조화가 시도되었기 때문이다. 이 현상은 파이프라인의 구조적 결함이 아니라 초기화 타이밍의 특성으로, 실제 운용 환경에서는 시작 후 일정 시간 대기 를 통해 회피할 수 있다. 본 연구에서는 TF 버퍼가 안정화되는 데 약 2초의 대기 시간이 필요한 것으로 확인되었으며, 이를 쿨다운 매커니즘으로 구현하여 적용하였다.

6. Comprehensive Discussion

전체 실험 조건에 대한 종합 비교를 Fig. 8에 나타내었다.

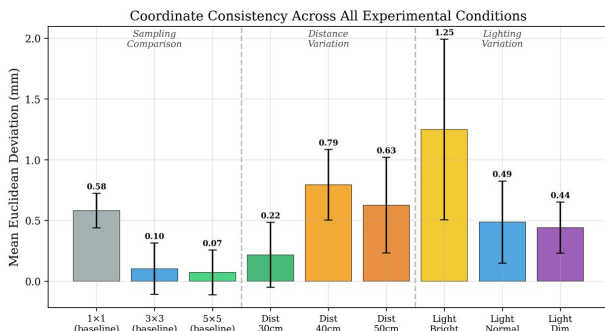


Fig. 8. Coordinate consistency across all experimental conditions

모든 실험 조건에서 좌표 변환 결과는 1.3mm 이하의 평균 유클리드 변동을 보였으며, 특히 최적 조건(5x5 샘플링)에서는 0.072mm의 변동을 달성하였다.

Shapiro-Wilk 정규성 검정 결과, 모든 조건에서 좌표 분포가 정규 분포를 따르지 않는 것으로 나타났다($p < 0.05$). 이는 깊이 센서의 양자화 특성에 기인하는 것으로, 깊이 값이 연속적이지 않고 이산적인 해상도 단위로 출력되기 때문이다. 이에 따라 본 연구에서는 비모수 통계 검정을 적용하여 조건 간 차이의 통계적 유의성을 검증하였다.

실험 결과를 종합하면 세 가지 주요 발견이 도출된다. 첫째, 깊이 샘플링 윈도우 크기는 좌표 변환 일관성에 가장 큰 영향을 미치는 설계 파라미터로, 3x3 이상의 윈도우를 사용하면 단일 픽셀 대비 5배 이상의 일관성 개선을 달성할 수 있다. 둘째, 카메라-대상 거리는 깊이 해상도를

통해 일관성에 영향을 미치며, 근거리 작업이 유리하다. 셋째, 조명 조건은 IR 기반 깊이 센서의 특성상 밝은 환경이 오히려 불리하게 작용하므로, 시스템 설계 시 이를 고려해야 한다.

V. Conclusions

본 논문에서는 ROS 기반 RGB-D 인식 파이프라인에서 깊이 샘플링 전략이 좌표 변환 일관성에 미치는 영향을 비교 분석하였다. 제안한 파이프라인은 RGB 영상 기반 인식, 윈도우 기반 깊이 샘플링을 포함한 3차원 좌표 산출, ROS TF 기반 좌표 변환, 그리고 제어 입력 좌표 정규화 단계로 구성된다.

총 270회(9개 조건 × 30회 반복)의 실험을 통해 다음의 결과를 확인하였다. 첫째, 깊이 샘플링 윈도우 크기가 좌표 변환 일관성에 통계적으로 유의한 영향을 미치며 (Kruskal-Wallis $H=71.59$, $p < 0.001$), 3x3 윈도우 중앙값 방식은 단일 픽셀 대비 5.7배, 5x5는 8.1배 일관성이 개선되었다. 둘째, 카메라-대상 거리가 증가할수록 좌표 변동이 증가하는 경향을 보이며(30cm: 0.217mm, 40cm: 0.794mm, 50cm: 0.626mm), 이는 깊이 해상도 변화에 기인한다. 셋째, 밝은 조명 환경에서 IR 센서 간섭으로 인한 변동 증가(1.248mm)가 확인되었으며, 이는 RGB-D 기반 파이프라인 설계 시 조명 조건을 고려해야 함을 시사한다.

본 연구는 좌표 변환 파이프라인의 내부 일관성 분석에 연구 범위를 한정하였으며, 실제 로봇의 수확 작업 수행이나 엔드이펙터 접촉 성공 여부에 대한 통합 성능 평가는 향후 연구 과제로 남긴다. 특히, 본 연구에서 확인된 좌표 변환 일관성이 실제 작업 성공률과 어떠한 상관관계를 보이는지를 검증하는 것은 중요한 후속 과제이다. 또한, 가림현상, 다중 객체 환경, 로봇 동작 중 진동 등 추가적인 환경 변수에 대한 분석도 후속 연구를 통해 보완할 필요가 있다.

그럼에도 불구하고, 깊이 샘플링 전략이라는 설계 파라미터가 좌표 변환 일관성에 유의한 영향을 미친다는 본 연구의 발견은 RGB-D 기반 로봇 시스템 설계 시 실용적인 지침을 제공할 수 있으며, 인식-제어 연계 시스템에서 좌표 변환 단계의 신뢰성을 사전에 검증하기 위한 기초 프레임워크로 활용될 수 있을 것으로 기대된다.

ACKNOWLEDGEMENT

This paper was supported by Suncheon National University Global University Project Fund in 2025.

REFERENCES

- [1] B. Siciliano and O. Khatib, "Springer Handbook of Robotics," Springer, 2nd Ed., 2016.
- [2] J. J. Craig, "Introduction to Robotics: Mechanics and Control," Pearson, 3rd Ed., 2019.
- [3] S. Izadi et al., "KinectFusion: Real-time 3D Reconstruction and Interaction Using a Moving Depth Camera," Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology (UIST), pp. 559-568, Oct. 2011. DOI: 10.1145/2047196.2047270
- [4] J. Sturm, N. Engelhard, F. Endres, W. Burgard, and D. Cremers, "A Benchmark for the Evaluation of RGB-D SLAM Systems," Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 573-580, Oct. 2012. DOI: 10.1109/IROS.2012.6385773
- [5] M. S. Ahn, H. Chae, D. Noh, H. Nam, and D. W. Hong, "Analysis and Noise Modeling of the Intel RealSense D435 for Mobile Robots," Proceedings of the 16th International Conference on Ubiquitous Robots (UR), pp. 707-711, Jun. 2019. DOI: 10.1109/URAI.2019.8768489
- [6] L. Rustler, V. Volprecht, and M. Hoffmann, "Empirical Comparison of Four Stereoscopic Depth Sensing Cameras for Robotics Applications," IEEE Access, Vol. 13, Apr. 2025. DOI: 10.1109/ACCESS.2025.3560810
- [7] X. Y. Han, J. S. Jin, M. J. Wang, W. Jiang, L. Gao, and L. P. Xiao, "A Review of Algorithms for Filtering the 3D Point Cloud," Signal Processing: Image Communication, Vol. 57, pp. 103-112, Sep. 2017. DOI: 10.1016/j.image.2017.05.009
- [8] M. Quigley et al., "ROS: An Open-Source Robot Operating System," Proceedings of the IEEE International Conference on Robotics and Automation (ICRA) Workshop on Open Source Software, Kobe, Japan, May 2009.
- [9] "tf2," ROS Wiki, <https://wiki.ros.org/tf2>
- [10] T. Yoshida, T. Kawahara, and T. Fukao, "Fruit Recognition Method for a Harvesting Robot with RGB-D Cameras," ROBOMECH Journal, Vol. 9, No. 15, 2022. DOI: 10.1186/s40648-022-00230-y
- [11] E. Ochoa et al., "Design and Field Evaluation of an End Effector for Robotic Strawberry Harvesting," Actuators, Vol. 14, No. 2, 42, 2025. DOI: 10.3390/act14020042
- [12] Z. He, Z. Liu, Z. Zhou et al., "Improving Picking Efficiency under Occlusion: Design, Development, and Field Evaluation of an Innovative Robotic Strawberry Harvester," Computers and Electronics in Agriculture, Vol. 237, 110684, Oct. 2025. DOI: 10.1016/j.compag.2025.110684
- [13] "D435i," librealsense2 Documentation, https://docs.ros.org/en/iron/p/librealsense2/user_docs/d435i.html

Authors



Kwang-Ho Yang received the B.S. degree in biology from Jeonbuk National University, Republic of Korea, in 2015, and the M.S. degree in smart agriculture from Suncheon National University, Republic of Korea, in

2023. Mr. Yang is currently a Ph.D. candidate in smart agriculture and a Researcher at the Low-Carbon Agriculture-Based Smart Distribution Research Center, Suncheon National University. His research interests include agricultural robotics, multi-modal image recognition, and artificial intelligence for smart farming systems.



Hyeon-O Choe received the B.S., M.S., and Ph.D. degrees in information and communication engineering from Suncheon National University, Republic of Korea, in 2017, 2019 and 2025, respectively.

Dr. Choe joined currently a Research Professor at the Low-Carbon Agriculture-Based Smart Distribution Research Center. His research interests include smart agriculture systems, digital twin-based environmental modeling, AI-based prognostics and health management (PHM) for agricultural equipment, sensor data analytics, and energy-efficient control algorithms for smart farms and vertical farming systems.