

농업 로봇의 3차원 위치 인식을 위한 스테레오 매칭 알고리즘 비교 연구

전성우*, 류진환*, 김만중*, 김경철*, 강상현*, 홍영기*, 이기범*

*국립농업과학원 농업공학부

e-mail: keywii@korea.kr

A Comparative Study of Stereo Matching Algorithms for 3D Perception of Agricultural Robots

Seong-Woo Jeon*, Jinhwan Ryu*, Manjung Kim*, Kyoung-Chul Kim*, Sang-Hyeon Kang*,
Youngki Hong*, Ki-Beom Lee*

*Department of Agricultural Engineering, National Institute of Agricultural Sciences, Jeonju, Korea

요약

본 논문에서는 스테레오 비전 알고리즘을 통해 딸기의 과실과 가늘고 텍스쳐가 거의 없는 줄기의 특성을 효과적으로 반영하는 스테레오 매칭 시스템의 성능을 비교하는 내용을 다룬다. 지역 최적화와 전역 최적화를 결합한 방식인 Semi-Global Block Matching 방식과 ZED SDK에서 제공하는 스테레오 매칭 방식, RUM과 AGCL을 결합하여 복잡한 이미지에서도 높은 정밀도의 시차 추정하는 CREStereo 알고리즘을 각각 시차 정보와 깊이 맵을 비교하였다. 비교 실험 결과 SGBM과 ZED Stereo 매칭 방식은 노이즈와 소실되는 데이터가 많은 반면, CREStereo 매칭 알고리즘은 가장자리에서 소실될 수 있는 시차 정보를 예측하는 능력을 갖추고 있어, 더 원활한 시차 계산을 가능하게 하고 시차 결과에서 정보 소실을 줄였다. 이러한 결과를 해석해 보면 CREStereo 알고리즘은 작물 모니터링과 3D 좌표 위치 정보를 효과적으로 취득하는 것이 가능할 것이다.

1. 서론

딸기(*Fragaria x ananassa*)는 풍부한 영양과 독특한 풍미로 인해 소비자들에게 높은 선호도를 보이는 작물이다[1]. 최근 토지 이용의 효율성을 극대화하고 연중 생산이 가능한 정밀 실내 농업 시스템(Precision Indoor Farming Systems, PIFS), 특히 인공광을 이용한 수직농장(Vertical Farm)이 딸기 재배의 유력한 대안으로 떠오르고 있다[2]. 정밀 농업 및 스마트 팜의 발전에 따라 식물의 특징을 정량적으로 기술하고 분석하는 것은 매우 중요한 과제이다. 특징 기술은 이미지로부터 식물의 고유한 특징이나 관심 지점을 추출하고 설명하는 과정으로, 식물 모니터링, 3D 복원, 객체 추적과 같은 핵심 응용 분야에 필수적이다. 특히 식물 높이, 꽃과 줄기의 밀도와 같은 형태학적 정보는 작물의 생육 상태를 진단하고 수확량을 예측하며, 재배 환경을 최적화하는 데 결정적인 데이터를 제공한다. 과거에는 이러한 데이터를 수동으로 측정하여 많은 시간과 노동력이 소요되었고, 측정의 정밀도와 일관성을 확보하기 어려웠다[3].

이러한 한계를 극복하기 위해 최근 농업 분야에서는 로봇 기술의 도입이 활발하게 이루어지고 있다. 수확, 방제, 시비 등 정밀한 작업을 수행하는 농업용 로봇은 주변 환경을 정확

하게 인식하고 대상 작물의 공간 정보를 파악하는 능력이 필수적이다. 이를 위해 최근에는 정밀한 3D 공간 정보를 제공하는 스테레오 비전(Stereo vision) 기술에 대한 연구가 활발하게 진행되고 있다. 스테레오 비전 기술은 LiDAR나 ToF와 같은 다른 3D 감지 기술에 비해 효율성이 높고, RGB 색상 정보와 깊이 정보를 동시에 취득할 수 있다는 장점이 있어 농업 재배 환경에 폭넓게 적용되고 있다[4]. 스테레오 비전 시스템에서 가장 중요한 과정 중 하나는 스테레오 매칭이다. 스테레오 매칭은 좌우 이미지에서 대응점을 찾아 시차를 계산하고, 이를 통해 깊이 정보를 추정하는 과정이다. 이 과정의 정확도와 효율성은 전체 시스템의 성능을 좌우하는 핵심 요소이다[5].

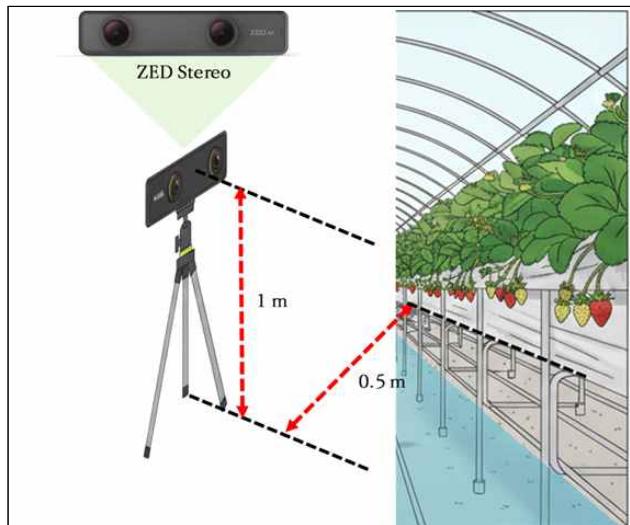
따라서 본 연구에서는 다양한 스테레오 매칭 알고리즘의 특성과 성능을 비교하는 과정을 진행했다. 특히 딸기의 줄기와 같이 가늘고 특정 패턴이 반복되거나 텍스쳐가 거의 없는 객체의 특성을 효과적으로 반영하는 스테레오 매칭 시스템의 성능 차이를 분석하고자 한다.

2. 실험 방법

2.1 이미지 데이터 취득

본 연구에서는 딸기의 3차원 데이터를 구성하기 위

해 실내 조명 환경에서의 모형 딸기 이미지 데이터셋과 외부 온실 환경에서의 딸기 이미지 데이터셋을 구축했다. 데이터 취득을 위해 국립농과학원 단동온실에서 ‘설향’ 품종의 딸기를 고설 방식으로 재배했으며, StereoLabs 사의 ZED mini 스테레오 비전 카메라를 활용하여 2025년 3월에 수집한 영상 데이터를 사용하였다.



[그림 1] 딸기 이미지 데이터를 얻기 위한 카메라 설치

2.2 스테레오 비전

스테레오 비전은 동일 장면을 서로 다른 시점에서 촬영한 두 개의 영상을 이용하여, 각 화소의 시차(disparity)를 계산하고 이를 기반으로 3차원 공간상의 깊이(depth) 정보를 추정하는 기술이다. 이는 인간의 양안 시각 원리와 유사하며, 정밀한 거리 측정을 가능하게 하여 로보틱스, 자율 주행, 정밀 농업 등 다양한 분야에서 핵심적으로 활용되고 있다. 스테레오 매칭은 이러한 깊이 추정을 위한 핵심 알고리즘으로, 좌·우 영상에서 대응점을 탐색하여 disparity map을 산출하는 과정이다.

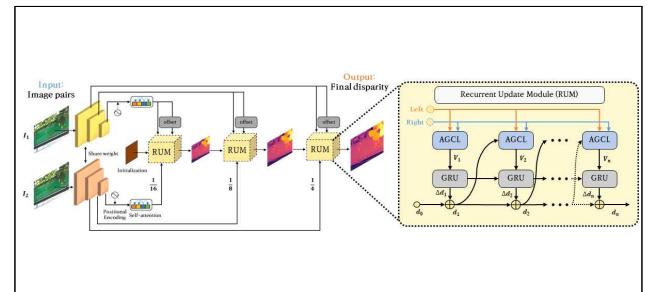
2.3 Semi-Global 블록매칭 방식

영상 분석에 널리 활용되어온 전통적 스테레오 매칭 알고리즘으로는 Semi-Global Block Matching (SGBM)과 같은 방법이 대표적이다. SGBM은 지역 최적화와 전역 최적화를 결합한 방식이다[6]. 각 픽셀에 대해 가능한 시차 값에 대한 매칭 비용을 계산하고, 이를 다 방향에서 누적하여 집계한다. 이 과정에서는 동적 프로그래밍 기법이 활용되며, 결과적으로 에너지 함수의 최적화를 통해 전역적인 일관성을 가진 시차 맵을 산출한다. 특히 시차 불연속성을 억제하기 위해 작은 시차 변화에 대해서는 소규모 폐널티(P_1)를, 큰 시차 변화에 대해서는 대규모

폐널티(P_2)를 부여하여 평활화를 유지한다.

2.4 CRESTereo 매칭

본 연구에서는 CRESTereo를 활용하여 농업 환경에서 3차원 영상 획득 및 분석을 수행하였다. 해당 네트워크는 RUM (Recurrent Update Module)과 AGCL (Adaptive Global Context Learning)을 결합하여, 복잡한 장면에서도 높은 정밀도의 시차 추정을 수행한다[7]. 특히, Deformable Search Window를 도입하여 이미지 내용에 따라 탐색 윈도우의 형태를 동적으로 변형시키는 오프셋을 학습한다. 이를 통해 가려짐(occlusion)이나 텍스쳐가 없고 세밀한 영역에서도 신뢰도가 높은 상관관계를 계산할 수 있으며, 이는 다단계(cascaded) 네트워크 구조를 통해 점진적이고 정교한 시차 보정이 가능하다.



[그림 2] CRESTereo 내부 구조

2.5 깊이 추정

깊이 추정을 위한 스테레오 비전 시스템의 파이프라인은 크게 Calibration, Rectification, Depth Estimation의 단계로 구성된다.

2.5.1 Calibration

Calibration 단계에서는 카메라의 내부 및 외부 파라미터를 산출한다. 내부 파라미터는 초점 거리(f_x, f_y), 주점 좌표(C_x, C_y), 그리고 방사 및 접선 왜곡계수(k_1, k_2, k_3, p_1, p_2)를 포함한다[8]. 이를 통해 렌즈와 센서의 구조적 특성으로 발생하는 영상 왜곡을 보정하여 왜곡 없는 영상을 생성할 수 있다. 외부 파라미터는 카메라의 좌표계를 실제 세계 좌표계에 정렬하기 위한 위치와 자세를 의미하며, 일반적으로 회전 행렬과 이동 벡터로 표현된다. ZED 카메라의 경우 공장 보정(factory calibration)된 파라미터를 SDK에서 불러와 사용할 수 있으며, 필요시 SDK의 self-calibration 기능을 통해 촬영 환경 변화에 따라 최적화된 보정을 수행할 수 있다.

2.5.2 Rectification

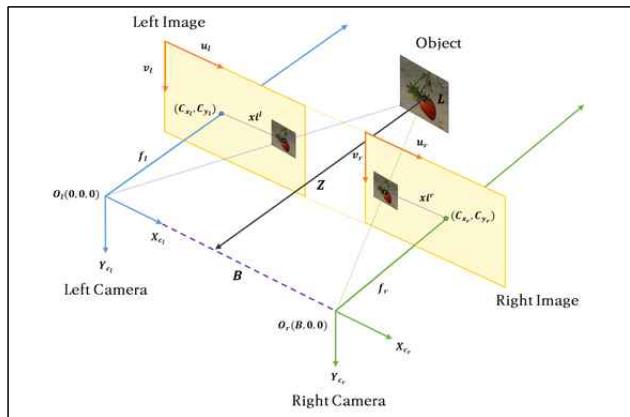
좌·우 카메라로부터 획득한 이미지 쌍을 기하학적으로 보정하여, 두 이미지의 주광선이 평행을 이루도록 정렬한다. 이를 통해 비정상적인 기하학적 차이를 제거하고, 두 이미지가 동일한 평면에서 수평 방향으로만 시차가 발생하도록 변환된다. 동일한 3차원 점은 좌·우 이미지에서 동일한 행에 위치하게 되며, 이는 이후에 수행되는 시차 계산의 복잡성을 줄이고 매칭 정확도를 높인다.

2.5.3 Depth estimation

깊이 추정은 보정된 좌·우 이미지에서 시차 맵을 산출한 뒤, 이를 카메라 기하학 모델을 이용하여 깊이 값으로 변환한다. 깊이는 삼각측량의 원리에 기반하며, 식은 다음과 같이 정의된다:

$$Z = \frac{f \cdot B}{d}$$

여기서, Z 는 특정 화소의 깊이, f 는 카메라의 초점거리(focal length), B 는 스테레오 카메라 렌즈 간의 기준 거리(baseline), d 는 해당 화소의 시차 값을 의미한다.



[그림 3] 삼각측량 기반 스테레오 비전의 원리

3. 결과 및 고찰

그림 4는 CRESTereo와 SGBM과 같은 전통적인 알고리즘 그리고 ZED SDK에서 제공하는 스테레오 매칭 결과를 비교하였다. CRESTereo 과정에서는 ETH3D 데이터셋 (Schöps et al., 2017)에서 사전 훈련된 가중치를 스테레오 매칭에 활용하였다[9].

그림 4에서 제시된 비교 결과를 바탕으로 다음과 같은 몇 가지 관찰 결과를 도출할 수 있었다.

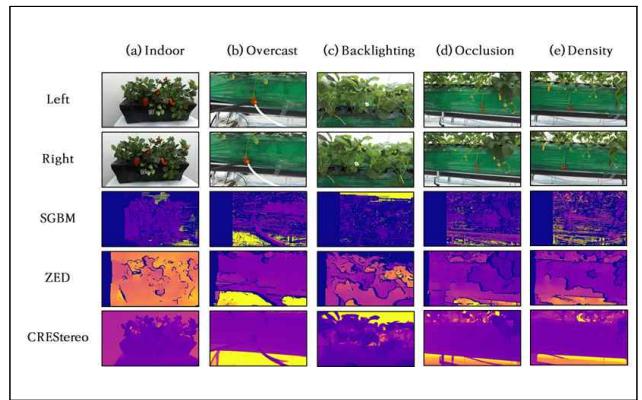
(1) 수직농장의 환경과 비슷한 인공광 조건 하에서, 그림 4(a)의 경우 SGBM으로 추출된 깊이 맵은 입력 영상들과 비교하였을 때 입력 영상들 속에 딸기 잎들을 모두 표현하였으나 노이즈가 많아 물체의 구분이 쉽지 않았다. 또한 시차 변환으로 인해 이미지의 상당 부분이 소실되는 것을 확인할 수 있었다.

(2) ZED SDK에서 제공하는 stereo 깊이 맵의 경우 영상 속의 물체의 구분이 어려웠으며, 노이즈가 많아 제대로된 깊이 값을 인식할 수 없었다.

(3) 그림 4(b)–(c)와 같이 불리한 광 조건에서도 CRESTereo는 딸기 잎과 줄기, 과실의 시차를 상당한 수준으로 효과적으로 유지하였다.

(4) 그림 4(d)처럼 잎에 의해 시야가 가려졌을 때, CRESTereo는 더 포괄적인 시차 정보를 제공하여 딸기 과실과 줄기의 이미지 정보의 누락이 더 적었다. 또한, 시차 노이즈가 감소하고 변동성이 최소화 되었다.

(5) CRESTereo 매칭은 스테레오 이미지의 가장자리에서 소실될 수 있는 시차 정보를 예측하는 능력을 갖추고 있어, 더 원활한 시차 계산을 가능하게 하고 시차 결과에서 정보 소실을 줄였다.



[그림 4] CRESTereo와 전통적인 스테레오 알고리즘 비교.

이러한 결과의 차이는 각 알고리즘이 가지는 구조적 특성과 딸기 재배 환경의 복잡성 사이의 상호작용에서 비롯된 것으로 고찰할 수 있다. SGBM과 같은 전통적인 블록 매칭 방식은 텍스쳐가 거의 없는 딸기 줄기나, 패턴이 반복되는 잎사귀 영역에서 대응점을 찾는데 본질적인 한계를 가진다. 이는 결과적으로 심한 노이즈와 정보 소실로 이어진다. ZED SDK 역시 범용적인 환경에서의 실시간성을 목표로 설계되었기 때문에, 이처럼 가늘고 복잡한 농업 환경에서는 정확한 깊이 정보를 제공하기 어려웠던 것으로 판단된다.

반면, CRESTereo는 Deformable Search Window와 같은 적응형 탐색 기법을 활용하여 딸기 줄기처럼 가늘거

나 앞에 가려진 객체에 대해서도 유연하게 대응점 탐색 영역을 조절할 수 있었다. 또한 다단계 구조를 통해 초기 추정된 시차를 점진적으로 정교하게 보정하는 과정을 거치므로, 최종적으로 노이즈가 적고 완성도가 높은 깊이 맵을 생성할 수 있었던 것으로 분석된다.

4. 결론

본 연구에서는 Semi-Global Block Matching (SGBM), CRESTereo 알고리즘의 이론적 배경과 작동 원리를 분석하고, 다양한 광 조건 하에서 스테레오 비전 알고리즘을 통해 딸기의 줄기와 잎, 과실 등의 깊이 맵의 성능을 평가하였다.

실험 결과, CRESTereo 알고리즘이 다른 방식에 비해 우수한 성능을 보였다. 특히, 시차 노이즈가 적고 깊이 차이를 잘 보여주며, 다양한 광 조건에서도 안정적인 성능을 보여주었다. 반면, SGBM 알고리즘은 전역 최적화 방법의 장점을 결합하여 보다 정확한 결과를 제공하였으나, 노이즈와 소실되는 부분이 많은 단점이 있었다.

앞으로의 연구에서는 더 다양한 데이터셋과 조건에서의 성능 평가를 통해 알고리즘의 적용 가능성을 확대할 필요가 있다. 또한 기존의 로봇이 정밀하게 접근하지 못하던 수확과 작물 모니터링 분야에서 딥러닝 기반의 스테레오 알고리즘을 통해 보다 정확한 깊이 맵을 적용하여 얻을 수 있을 것으로 판단된다.

5. 감사의 글

This research was supported by the Technology Innovation Program (RS-2024-00443366, “Agriculture dual-arm robot autonomy technology based on force posture composite learning”) funded by the Ministry of Trade, Industry and Energy, Korea Government.

참고문헌

- [1] Francesca Giampieri, Sara Tulipani, Josè M. Alvarez-Suarez, Josè L. Quiles, Brouno Mezzetti, Maurizio Battino, “The Strawberry: Composition, nutritional quality, and impact on human health”, *Nutrition*, 28(1), 9–19, 2012.
- [2] Ren Guoqiang, Wu Hangyu, Bao Anbo, Lin Tao, Ting Kuan-Chong, Ying Yibin, “Mobile robotics platform for strawberry temporal-spatial yield monitoring within precision indoor farming system”, *Frontiers in Plant Science*, 14, 2023.
- [3] Karim Md, Ahmed Shahriar, Reza Md Nasim, Lee Kyu-Ho, Jin Hongbin, Ali Mohmmad, Chung Sun-Ok, Sung Joonjea, “A review on stereo vision for feature characterization of upland crops and orchard fruit trees”, *Precision Agriculture Science and Technology*, 6(2), 104–122, 2024.
- [4] Wang Q, Wu D, Liu W, Lou M, Jiang H, Ying Y, Zhou M, “A High Quality Stereo Matching Dataset for Plant Reconstruction”, *Agriculture*, 13(2), 330, 2023.
- [5] 정유선, 김지수, 이덕우, “스테레오 매칭 알고리즘 성능 비교 연구: BM, SGBM, ELAS의 비교”, *한국산학기술학회논문지*, 제 25권 11호, pp 667–673, 11월, 2024년.
- [6] E. Mezenceva, S. Malakhov, “The Study of the Semi-Global Block Matching Algorithm Implementing Parallel Calculation with GPU”, *2021 IEEE 8th International Conference on Problems of Infocommunications, Science and Technology (PIC S&T)*, 2021.
- [7] Jiankun Li, Peisen Wang, Pengfei Xiong, Tao Cai, Ziwei Yan, Lei Yang, Jiangyu Liu, Haoqiang Fan, Shuaicheng Liu, “Practical Stereo Matching via Cascaded Recurrent Network with Adaptive Correlation”, *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 16263–16272, 2022.
- [8] Ortiz Luis Enrique, Elizabeth V. Cabrera, Luiz M. Gonçalves, “Depth data error modeling of the ZED 3D vision sensor from stereolabs” *ELCVIA. Electronic letters on computer vision and image analysis*, 17(1), 0001–15, 2018.
- [9] Schöps T, Schönberger JL, Galliani S, Sattler T, Schindler K, Pollefeys M, “A multi-view stereo benchmark with high-resolution images and multi-camera videos”, *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.