(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2023년08월30일
(11) 등록번호 10-2572571
(24) 등록일자 2023년08월25일

(51) 국제특허분류(Int. Cl.)
G06T 7/70 (2017.01) *A01D 46/30* (2006.01)
B25J 9/16 (2006.01)
(52) CPC특허분류
G06T 7/70 (2017.01)
A01D 46/30 (2013.01)
(21) 출원번호 10-2022-0151525
(22) 출원일자 2022년11월14일
심사청구일자 2022년11월14일
(56) 선행기술조사문헌
Jizhan Liu외 2인, "Experimental and theoretical analysis of fruit plucking patterns for robotic tomato harvesting", ELSEVIER Computers and Electronics in Agriculture vol.173, (2020.06.30.) 1부.*
M. Benavides외 3인, "Automatic tomato and peduncle location system based on computer vision for use in robotized harvesting", MDPI, (2020.08.25.) 1부.*
Zhe Cao외 3인, "Realtime multi-person 2D pose estimation using part Affinity Fields", CVPR 2017, (2017.07.26.) 1부.*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
충남대학교산학협력단
대전광역시 유성구 대학로 99 (궁동, 충남대학교)
(72) 발명자
이대현
세종특별자치시 새롬남로 18 새뜸마을1단지
(74) 대리인
이충한

전체 청구항 수 : 총 1 항 심사관 : 박금옥

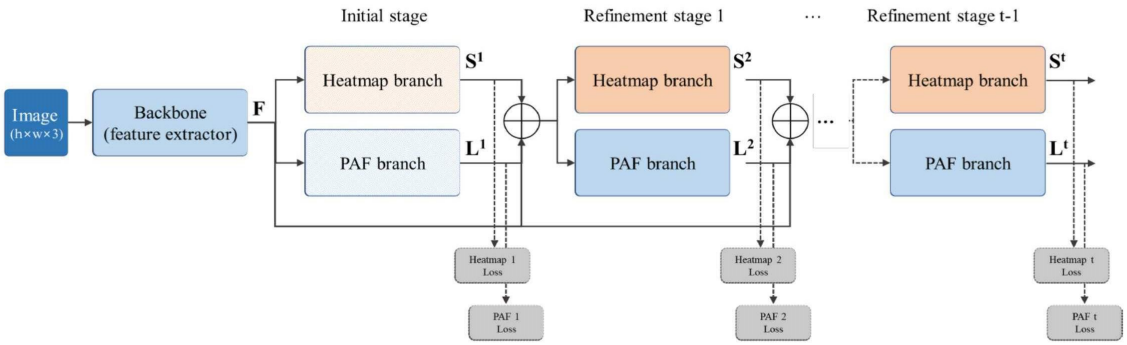
(54) 발명의 명칭 **로봇 수확을 위한 과일 줄기 자세 인식 시스템**

(57) 요 약

본 발명은 로봇 수확을 위한 과일 줄기 자세 인식 시스템에 관한 것으로 모든 과일-줄기의 주요부위들을 우선 검출 후 이를 상관관계에 따라 연결해주는 방식으로 개체수에 상관없이 일관적인 검출 및 동시다발적으로 다양한 과일-줄기 쌍의 자세 검출이 가능한 로봇 수확을 위한 과일 줄기 자세 인식 시스템에 관한 것이다. 구체적으로

(뒷면에 계속)

대 표 도



본 발명의 일양태에 따르면, 로봇 수확을 위한 과일 줄기 자세 인식 시스템이 제공되고, 이 시스템은, 과일의 영상을 취득하기 위한 영상 취득부; 및 영상 취득부로부터 취득된 영상으로부터 과일-줄기 골격 정보를 추정하기 위한 인공지능기반의 과일-줄기 골격 추정부를 포함하고, 상기 과일-줄기 골격 정보는, 키포인트1로서 과일 중심부(fruit center); 키포인트2로서 과일과 줄기가 연결되는 꽃받침(calyx); 키포인트3으로서 소화경(pedicel) 중심에 위치한 탈리대(abscission zone); 키포인트4로서 소화경과 화방(peduncle)의 가지점(branch point)의 4개의 키포인트를 이용하여 구성될 수 있다.

(52) CPC특허분류

B25J 9/1664 (2013.01)

G06T 2207/20044 (2013.01)

G06T 2207/20081 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1545026083
과제번호	421031042HD070
부처명	농림축산식품부
과제관리(전문)기관명	농림식품기술기획평가원
연구사업명	스마트팜다부처패키지혁신기술개발사업
연구과제명	작업 대상체의 3차원 위치 및 방향각 추출기술개발
기 여 율	1/1
과제수행기관명	충남대학교 산학협력단
연구기간	2021.04.07 ~ 2024.12.31

공지예외적용 : 있음

명 세 서

청구범위

청구항 1

로봇 수확을 위한 과일 줄기 자세 인식 시스템에 있어서,

과일의 영상을 취득하기 위한 영상 취득부;

영상 취득부로부터 취득된 영상으로부터 과일-줄기 골격 정보를 추정하기 위한 인공지능기반의 과일-줄기 골격 추정부를 포함하고,

상기 과일-줄기 골격 정보는,

키포인트1로서 과일 중심부(fruit center); 키포인트2로서 과일과 줄기가 연결되는 꽃받침(calyx); 키포인트3으로서 소화경(pedicel) 중심에 위치한 탈리대(abscission zone); 및 키포인트4로서 소화경과 화방(peduncle)의 가지점(branch point)의 4개의 키포인트를 이용하여 구성되고,

상기 과일-줄기 골격 정보는 과일-줄기 쌍으로서, 키포인트1과 키포인트2를 연결하는 라인; 키포인트2와 키포인트3을 연결하는 라인; 및 키포인트3과 키포인트4를 연결하는 라인;를 포함하고,

상기 과일-줄기 골격 추정부는 하나의 영상 내 과일-줄기 쌍이 복수개 존재하는 경우 PAF (part affininty field)를 이용하여 각 개체별 키포인트 1-2-3-4 연결강도를 추정하도록 학습되데,

상기 과일-줄기 골격 추정부에서의 토마토 포즈의 추정은 신뢰 맵에서 후보 키포인트 위치를 획득하고 PAF를 기반으로 이들 간의 연결을 찾는 방식으로 수행되고, 예측된 신뢰 맵에서 피크 포인트를 탐지하고 NMS(Non-Maximum Suppression)를 수행하여 노이즈를 제거하여 후보 키포인트 위치를 추정하고, 여러 토마토의 키포인트가 하나의 이미지에 함께 있기 때문에 PAF를 사용하여 각 토마토의 연결을 포함하여 전체 포즈를 추정하고, 부분에 포함된 두 개의 키포인트 후보의 연관 점수는 다음의 식(7)과 같이 두 후보 위치 dj1, dj2를 연결하는 선분을 따라 해당 PAF의 적분을 계산하여 구하고,

$$E = \int_{u=0}^{u=1} \mathbf{L}_c(\mathbf{p}(u)) \cdot \frac{\mathbf{d}_{j_2} - \mathbf{d}_{j_1}}{\|\mathbf{d}_{j_2} - \mathbf{d}_{j_1}\|_2} du \qquad (7)$$

$$\mathbf{p}(u) = (1-u)\mathbf{d}_{j_1} + u\mathbf{d}_{j_2}$$

여기서 p(u)는 두 개의 키 포인트 dj1 및 dj2의 보간이고,

각 부분 유형에 대한 연결 가능한 모든 키포인트 후보에 대해 연관 점수를 얻어 NMS를 수행하고, 연관 점수가 높은 순서대로 부분을 연결하고, 한번 연결이 확인된 키포인트 후보는 다른 키포인트와의 연결 후보에서 제외하고, 토마토 포즈의 마지막 부분인 근위 소화경은 NMS를 사용하지 않고 가장 높은 연관 점수를 가진 가지 키포인트 4을 연결하여 여러 토마토의 부분이 공유할 수 있도록 구성되고,

상기 개체별 키포인트 중 키포인트4인 가지점은 개체별 공유가능하도록 학습되고, 추정된 과일-줄기 골격 정보를 수확 로봇으로 전송하기 위한 추정 정보 전송부를 더 포함하고, 수확 로봇은 전송된 과일-줄기 골격 정보에 기반하여 수확 로봇의 그리핑 위치를 결정하도록 구성된 것을 특징으로 하는

로봇 수확을 위한 과일 줄기 자세 인식 시스템.

청구항 2

삭제

청구항 3

삭제

청구항 4

삭제

청구항 5

삭제

발명의 설명

기 술 분 야

[0001] 본 발명은 로봇 수확을 위한 과일 줄기 자세 인식 시스템에 관한 것으로 모든 과일-줄기의 주요부위들을 우선 검출 후 이를 상관관계에 따라 연결해주는 방식으로 개체수에 상관없이 일관적인 검출 및 동시다발적으로 다양한 과일-줄기 쌍의 자세 검출이 가능한 로봇 수확을 위한 과일 줄기 자세 인식 시스템에 관한 것이다.

배 경 기 술

[0003] 국내 농업은 지속적인 고령화 및 농업인 감소로 인한 인력부족 뿐만 아니라 기후 변화, 경영 악화, 경지 감소 등에 따른 식량 안보 문제 해결을 필요로 하고 있어 향후 자동화-로봇화 농업기계 기술의 중요성은 더욱 커질 것으로 예상된다.

[0005] 지능-정보 융합 핵심기술인 농업용 로봇의 경우 시장 규모는 `13년 9억불에서 `20년 163억불까지 급성장할 것으로 예상되고 있으며, 국내 농업용 로봇은 `30년까지 4조 2천억원 규모로 성장할 것으로 추정되고 있다.

[0007] 농업용 로봇은 예찰, 파종, 접목, 수확 등과 같은 농업 생산 작업의 자동화, 무인화를 실현할 수 있으며 이를 통해 노동력 부족 문제 해결 및 국가 농업경쟁력 확보가 가능하다.

[0009] 특히, 수확로봇은 농작업 중 작업 과정과 목표의 정의가 비교적 명확하여 로봇화의 가능성이 가장 높으며, 해외의 경우 파프리카 등 과채류를 대상으로 수확로봇을 개발하여 제품화 수준까지 기술을 보유한 상태이다.

[0011] 국내에서도 시설원예 중심으로 농업용 수확로봇이 연구개발되고 있으며, 기존 기술력을 활용하여 대상물 인식 및 수확을 위한 매니플레이팅 등은 기술력을 점차 확보하고 있으나 실제 수확을 위한 대상체 인식의 실용적인 기술 확보가 어려운 상태이다.

[0013] 수확을 위한 인식기술은 대부분 영상 기반으로 진행되며 주로 수확 목표인 과일이나 과일이 매달려있는 줄기 인식하여 로봇팔(매니플레이터+그리퍼)에 위치 및 자세 정보를 제공하게 된다.

[0015] 현재까지 개발되는 대부분의 기술들은 대상 과일이 선정되었다는 가정 하에 하나의 개체에 대한 검출을 수행하거나 전체 영상에 존재하는 과일 하나하나에 대해 인식을 개별적으로 수행하는데, 전자의 경우 해당 과일 혹은 줄기가 생육상태나 차폐 등에 따라 수확이 불가능한 경우일 수 있고 후자의 경우 개체수에 따라 연산량이 늘어나 효율을 저하시킬 수 있다.

[0017] 따라서 전체 영상에 대한 개체(과일-줄기)들의 동시다발적인 인식으로 여러 개체의 위치 및 자세 정보를 종합적으로 확인할 수 있어야 그 중 수확 가능성이 있는 개체의 선정이 가능하며, 개체수에 따른 연산량의 변화가 크지 않아야 일관된 성능을 유지할 수 있지만, 현재까지 그와 같은 과일 수확 시스템은 알려지지 않았다.

선행기술문헌

특허문헌

[0019] (특허문헌 0001) 한국등록특허 10-1405858(2014년06월03일)

발명의 내용

해결하려는 과제

[0020] 본 발명은 전술한 문제점에 기반하여 종합적인 과일-줄기의 위치와 자세 정보 제공함으로써 효율적이고 실용적

인 수확 대상체 인식이 가능한 로봇 수확을 위한 토마토 줄기 자세 인식 시스템을 제공하는 것을 목적으로 한다.

과제의 해결 수단

[0022] 전술한 과제를 해결하기 위해, 본 발명의 일양태에 따르면, 로봇 수확을 위한 과일 줄기 자세 인식 시스템이 제공되고, 이 시스템은,

[0023] 과일의 영상을 취득하기 위한 영상 취득부; 및 영상 취득부로부터 취득된 영상으로부터 과일-줄기 골격 정보를 추정하기 위한 인공지능기반의 과일-줄기 골격 추정부를 포함하고, 상기 과일-줄기 골격 정보는, 키포인트1로서 과일 중심부(fruit center); 키포인트2로서 과일과 줄기가 연결되는 꽃받침(calyx); 키포인트3으로서 소화경(pedicel) 중심에 위치한 탈리대(abscission zone); 키포인트4로서 소화경과 화방(peduncle)의 가지점(branch point)의 4개의 키포인트를 이용하여 구성될 수 있다.

[0025] 전술한 양태에서, 상기 과일-줄기 골격 정보는 과일-줄기 쌍으로서, 키포인트1과 키포인트2를 연결하는 라인; 키포인트2와 키포인트3을 연결하는 라인; 및 키포인트3과 키포인트4를 연결하는 라인;를 포함한다.

[0027] 또한 전술한 어느 하나의 양태에서, 과일-줄기 골격 추정부는 하나의 영상 내 과일-줄기 쌍이 복수개 존재하는 경우 PAF (part affininty field)를 이용하여 각 개체별 키포인트 1-2-3-4 연결강도를 추정하정하도록 학습된다.

[0029] 또한 전술한 어느 하나의 양태에서, 개체별 키포인트 중 키포인트4인 가지점은 개체별 공유가능하도록 학습된다.

[0031] 또한 전술한 어느 하나의 양태에서, 추정된 과일-줄기 골격 정보를 수확 로봇으로 전송하기 위한 추정 정보 전송부를 더 포함하고, 수확 로봇은 전송된 과일-줄기 골격 정보에 기반하여 수확 로봇의 그리핑 위치를 결정하도록 구성된다.

발명의 효과

[0033] 본 발명에 따르면 본 발명은 전술한 문제점에 기반하여 종합적인 과일-줄기의 위치와 자세 정보 제공하고, 효율적이고 실용적인 수확 대상체 인식이 가능한 로봇 수확을 위한 토마토 줄기 자세 인식 시스템을 제공할 수 있다.

도면의 간단한 설명

[0035] 도 1은 토마토 열매 맺는 시스템의 구조와 키포인트를 결정하기 위한 주요 구성 요소를 나타내는 도면;

도 2는 토마토 열매 맺는 시스템의 포즈 골격의 예를 나타내는 도면 - 4개의 키포인트와 순차적인 페어링을 사용하는 토마토 과일 베어링 시스템의 포즈를 나타내고, 여기서 오른쪽 및 왼쪽 부분은 각각 단일 및 다중 토마토의 골격용이고, 키포인트와 페어는 열매중심(key 1)은 파란색, 꽃받침은 초록색, 이탈부는 빨간색, 가지점은 노란색으로 색상으로 표현하였으며, 쌍의 색상은 시작 키포인트와 동일함 - ;

도 3은 OpenPose 아키텍처를 네트워크 아키텍처를 나타내는 도면;

도 4는 멀티 토마토 포즈 추정의 대표적인 결과를 나타내는 도면 - 토마토 열매 맺는 시스템에 대한 대표적인 포즈 추정 및 30개 샘플의 결과가 6행 5열로 표현됨, 결과는 대상 개체의 수에 따라 행별로 정렬되며 각 행은 5개의 샘플을 나타냄, 완전히 익었는지 여부에 관계없이 모든 토마토에서 포즈가 감지됨 - ;

도 5는 토마토 개체 간의 가지 점 공유에 대한 근위 소화경부를 추정하기 위해 NMS 방법을 적용하지 않은 효과의 결과를 나타내는 도면 - 토마토 포즈 추정을 위한 가지 점 공유 효과에 대한 결과. 왼쪽과 오른쪽 그림은 각각 가지 점 비공유 및 공유의 대표적인 예를 나타내고, 각 이미지의 흰색 화살표는 가지 점 공유를 통해 감지되지 않은 소화경이 감지된 연결 부분을 가르킴 - ;

도 6은 정규화된 거리에 따른 검출 정확도를 나타내는 도면 - 정규화된 거리에 따른 4개의 키포인트(왼쪽)와 3개의 쌍(오른쪽) 간의 PDK 곡선 비교함 - ;

도 7은 길이와 각도 오차로 표현되는 추정된 포즈의 성능을 나타내는 히스토그램 - 추정된 쌍의 길이(왼쪽) 및 각도(오른쪽) 오류에 대한 확률 분포와 행은 위에서부터 토마토 열매, 원위 소화경 및 근위 소화경에 대한 것임

- ;

도 8은 본 발명에 따른 로봇 수확을 위한 과일 줄기 자세 인식 시스템의 일례를 나타내는 도면이다.

발명을 실시하기 위한 구체적인 내용

[0036] 본 발명의 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 상세하게 후술되는 실시예를 참조하면 명확해질 것이다. 그러나, 본 발명은 이하에서 개시되는 실시예에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 것이다. 본 명세서에서 본 실시예는 본 발명의 개시가 완전하도록 하며, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이다. 그리고 본 발명은 청구항의 범주에 의해 정의될 뿐이다. 따라서, 몇몇 실시예들에서, 잘 알려진 구성 요소, 잘 알려진 동작 및 잘 알려진 기술들은 본 발명이 모호하게 해석되는 것을 피하기 위하여 구체적으로 설명되지 않는다.

[0038] 본 명세서에서 사용되는 모든 용어(기술 및 과학적 용어를 포함)는 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 공통적으로 이해될 수 있는 의미로 사용될 수 있을 것이다. 또 일반적으로 사용되는 사전에 정의되어 있는 용어들은 정의되어 있지 않은 한 이상적으로 또는 과도하게 해석되지 않는다.

[0040] 이하 첨부된 도면을 참조하여 본 발명의 실시예에 대해 보다 상세하게 설명하도록 한다.

[0042] 1. 서론

[0043] 농업의 상시적인 노동력 부족은 농업공학 분야의 로봇 농업에 대한 요구를 요구하고 있다. 과일 수확은 그와 관련된 파지 과정이 산업용 로봇의 일반적인 파지 시스템(물체 인식, 유지 및 이동)과 유사하기 때문에 인간 대신 로봇이 수행할 가능성이 높다.

[0045] 실제로 수확 로봇은 다양한 목적으로 널리 개발되어 왔으며 일반적으로 과일 감지, 성숙도 결정, 장애물 회피 및 조작자의 궤적 계획을 포함한 여러 시퀀스를 통합하여 구현된다.

[0047] 특히 과일 감지는 성공적인 로봇 수확을 위한 최우선 요구 사항이며 위치, 방향 및 모양과 같은 수확 대상 과일에 대한 기하학적 정보를 제공할 수 있어야 한다.

[0049] 과일을 감지하기 위해 깊이 맵이 있는 컬러 이미지는 포인트 클라우드라고 하는 3차원(3D) 기하 정보뿐만 아니라 시각적 표현을 제공하는 데 자주 사용된다.

[0051] 베레나(Barnea)는 대상 과일의 색상과 모양 변화를 극복하기 위해 시각적 하이라이트 기반 달콤한 고추 감지를 제안했으며 최대 0.55mAP로 감지를 수행할 수 있다고 보고했다. 접근 방식은 실제 수확 로봇 플랫폼을 사용하여 평가되었다. 렌허트(Lehnert) 는 또한 엔드 이펙터의 수확 자세를 결정하기 위해 달콤한 고추의 검출을 연구했다. 이 접근법은 작물의 6-DOF(자유도) 자세를 추정할 수 있다. 성능 테스트는 제안된 알고리즘을 사용하는 로봇 팔이 평균 80%의 성공률로 달콤한 고추 몸체를 잡을 수 있음을 나타냈다. 이러한 과제는 심층 신경망(DNN)과 같은 데이터 기반 접근 방식의 급속한 발전으로 적극적으로 추구되고 있다. 지난 10년 동안 이미지 인식에 널리 사용되며 환경에 강력한 기능을 추출할 수 있는 DNN 유형인 CNN(컨볼루션 신경망)의 개발로 자동 시각적 감지가 가능해졌다.

[0053] 딥 러닝은 수확 로봇을 사용한 과일 감지에 상당한 잠재력을 가지고 있다. 여러 연구에서 딥 러닝이 열매를 파악하는 데 유용한 정보를 제공하는 놀라운 성능을 보여주었다. 종래 연구에서는 로봇 팔을 제어하기 위해 과일 포즈를 추정한 3D sphere Hough Transformation(3D-SHT)이 있는 DNN을 사용하여 사과 분할 및 수확 환경 모델링을 제시했다. 이 연구에서는 각각 0.87과 0.86의 F1-score와 IoU(intersection over union)로 상당한 성능을 보여주었다. 또 다른 연구에서는 쌍안 이미지를 사용하여 마스크 R-CNN을 기반으로 포즈로 포도송이를 감지했다. 그들은 그들의 방법이 수확 로봇에 적용하기에 충분히 정확하게 표적을 추출할 수 있다고 언급했다. 앞서 언급한 연구 외에도 사과 과수원의 여러 과일과 한 상자에 들어 있는 천연 과일(사과와 바나나)의 검출 및 포장을 위한 40개 과일의 위치 및 방향 추정이 실제 시스템에 사용하기 위해 수행되었다.

[0055] 일반적으로 수확을 위해 과일을 분리하는 방법에는 따기와 자르기(절단)의 두 가지가 있다. 따기 방식은 열매체를 잡고 비틀거나 구부리거나 잡아당겨 열매를 떼어내는 방식이다. 따기 방식은 열매를 맺는 가지나 줄기의 절단 지점을 추가로 감지해야 하는 절단 방식보다 간단한다. 그러나 몸체를 잡으면 과실 표면이 손상되어 과실 품질이 저하될 위험이 있다. 또한, 따기 방식의 간접적인 과실 박리로 인해 과실의 형태가 다양할 뿐만 아니라 가지의 물리적 특성도 고려하여야 한다. 따라서 따기 방식에는 복잡한 그립 메커니즘이 필요한다. 절단 방식은 열매가 있는 가지를 직접 절단하여 열매를 분리한다. 절단 지점을 정확하게 감지하면 과일 품종에 대한 그리퍼의

의존도를 줄일 수 있다. 이러한 맥락에서 최근 절단형 로봇 수확을 위한 가지 감지가 연구되고 있다.

[0057] 예를 들어 종래의 연구에서는 토마토를 자동으로 수확하는 꽃자루 감지 방법을 제안했다. 이 방법은 거친 토마토 현지화와 여러 꽃자루의 픽셀 수준 세분화로 구성된다. 이 연구에서는 접근 방식이 각도 오류에 대해 최대 5° 로 토마토 꽃자루 자세를 추정할 수 있다고 보고했다.

[0059] 또한 벚나무 가지의 떨림 위치 파악, 연속 로봇 따기에 사용하기 위한 열매 맺는 가지 감지, 고추 꽃자루 추정을 통해 엔드이펙터 포즈 결정, 토마토 위치 파악 등의 괄목할 만한 실용적인 연구가 진행되었다. 가지 감지는 실제 환경에서 사용하기 위한 상당한 성과를 보여주었지만 수확 로봇에 적용하기에는 여전히 한계가 있다.

[0061] 대부분의 경우 가지 감지 방식은 기하학적, 형태학적 요인 및 선 맞춤을 포함한 다양한 기능을 사용하여 과일 감지 및 해당 가지를 추정하는 여러 단계로 구성된다.

[0063] 기존의 방법은 대상 토마토를 먼저 감지한 다음 수확할 가지의 절단 지점을 추정한다. 그러나 실제 환경에서는 대상의 절단 지점이 자세나 다른 토마토와 가려져 시각적 추정이 어려울 수 있다.

[0065] 대상 과일이 감지되어도 절단 지점 감지에 실패하면 로봇은 절단 지점 감지가 가능해질 때까지 토마토 감지 및 절단 지점 추정을 포함한 이러한 일련의 작업을 반복적으로 수행해야 한다.

[0067] 하향식 방식이라고 하는 이 시퀀스는 물체 검출 결과를 이용한 N배의 목표 추정이 필요하기 때문에 일반적으로 다수의 물체를 시각적으로 검출하기 위한 처리 시간을 소모한다. 이 방법은 수확 시간을 늘리고 상당한 컴퓨팅 성능을 필요로 하기 때문에 비효율적이다.

[0069] 딥 러닝 응용 분야에서 중요한 관심 분야인 포즈 추정은 이미지 프레임에서 물체의 위치와 방향을 감지하는데, 이는 활동 인식, 모션 추적, 증강 현실과 같은 다양한 목적을 위한 컴퓨터 비전 분야의 일반적인 문제이다.

[0071] 포즈 추정은 객체의 키포인트 간의 연결을 감지하여 수행된다. 예를 들어, 인간 포즈 추정은 다양한 인간 관절(예: 머리, 목, 손목, 팔꿈치)을 키포인트로 사용하고 이들의 연결을 추정한다.

[0073] 이 도전은 먼저 각 사람을 감지하고 사람의 관절을 추정하는 하향식 방법을 사용하여 수행되었다. 이에 반해 앞서 언급한 한계를 극복하기 위해 최근에는 영상에서 모든 관절을 먼저 찾아 사람의 자세를 추정하는 상향식(bottom-up) 방식이 제안되고 있다.

[0075] 이 방법은 검출 대상의 개수에 관계없이 비교적 일정한 연산량이 요구되는 장점이 있다. 동시에 여러 객체를 처리했지만 잘 수행되었다. 따라서 상향식 방법은 눈에 보이는 과일에서 어떤 과일을 수확해야 하는지 결정하는 데 더 효과적일 수 있다.

[0077] 최근 농업 분야에서 포즈 추정을 이용하여 동물의 행동을 자동으로 모니터링하여 소, 닭과 같은 동물의 건강 및 이상 여부를 판단하는 데 유용한다. 이러한 접근 방식은 단일 개체에 대해 잘 작동했다. 그러나 로봇 수확은 여러 과일 맺는 시스템의 자세를 추정하는 데 더 어려운 작업이다. 따라서, 본 발명는 과실 맺는 시스템의 2차원 자세를 추정하는 것을 목적으로 하였다.

[0079] 가장 경제적으로 중요한 작물 중 하나인 토마토를 대상작물로 선정하였다. 여러 토마토와 그 가지에 대한 자세 추정에 기반한 상향식 방법은 수확할 대상 토마토를 직관적으로 결정하는 데 유용한 기하학적 정보를 제공할 수 있다.

[0081] 사람 포즈 추정을 위해 제안된 OpenPose를 백본 아키텍처로 사용하고, 인간의 관절과 달리 과일 사이에서 가지를 공유할 수 있는 토마토 베어링 시스템의 포즈를 추정하기 위한 프레임워크를 설정했다.

[0083] 종단 간 프로세스를 사용하여 여러 토마토 과일과 관련 가지의 자세 추정을 조사한 연구는 아직까지 알려지지 않았다. 본 발명에 따른 방법은 전체 장면에서 여러 토마토 과일과 그 가지를 동시에 추정하는 처리 시간을 줄일 수 있다.

[0085] 2. 재료 및 방법

[0086] 2.1. 토마토 열매 맺는 시스템 설명의 포즈

[0088] 토마토 열매 맺는 시스템은 도 1과 같이 정의되며, 줄기와 열매 사이에 생물학적으로 구별되는 꽃자루(peduncle), 소화경(pedicel), 탈리대(abscission zone), 꽃받침(calyx)과 같은 여러 부분을 포함한다.

[0090] 꽃자루는 원줄기에서 나온 가지로서 하나 또는 여러 개의 작은 꽃자루가 있고 가지에 개별 열매가 달려 있다.

소화경은 해부학적으로 탈리대에 따라 원위(distal) 소화경과 근위(proximal) 소화경으로 나눌 수 있으며, 이 소화경에서 잘 익은 열매는 생리학적으로 분리 과정에 의해 박리된다.

[0092] 탈리대는 꽃받침과 원위 소화경을 모두 유지하면서 과실 분리를 용이하게 하기 때문에 기계적 수확에 중요한다. 여기서 꽃받침은 열매 꼭대기에 꽃 모양의 잎이 달린 꽃받침이라고도 한다.

[0094] 이러한 맥락에서, 자동 수확을 위한 소화경을 감지하기 위한 가장자리로 탈리대(albscission zone)을 사용하였고, 로봇의 수확을 보다 쉽게 하기 위해 기계적 특성을 분석하였다. 또한 주변과 생김새가 달라 식별이 용이하기 때문에 과수계의 자세 추정 등 시각적 인지의 키포인트로 활용되는 장점이 있다.

[0096] 따라서 과일과 탈리대를 가진 작은 꽃자루(소화경)로 구성된 토마토 열매 맺는 시스템의 자세 추정은 자동 수확을 위한 포괄적인 기하학적 정보를 효과적으로 제공할 수 있다.

[0098] 본 발명에서 토마토 열매 맺는 시스템의 자세는 1-과실 중심, 2-꽃받침, 3-탈리대 및 4-가지 점(branch point)의 순서로 된 4개의 키포인트를 사용하여 설정되었다. 핵심 포인트는 해당 영역의 중심점을 사용하여 시각적으로 결정되었다.

[0100] 키포인트 1은 열매의 모양이 원과 일직선일 때 원의 중심이고, 키포인트 2는 꽃받침과 소화경이 연결되는 부분의 중심이다. 키포인트 3은 소화경보다 상대적으로 더 볼록한 띠인 탈리대의 중심이다. 키포인트 4는 본 발명에서 꽃자루와 소화경(작은 꽃자루)를 연결하는 지점으로 정의한 가지 점이다.

[0102] 두 개의 연속된 키포인트는 연결되어 키포인트 간의 기하학적 관계를 나타내는데, 이는 본 발명에서 부분(part)이라고 하며, 토마토 열매 맺는 시스템의 자세를 결정하는 요소이다.

[0104] 키포인트 1-2, 2-3, 3-4를 각각 연결하는 선과 과실축(part 1), 원위 소화경(part 2), 근위 소화경(part 3)을 연결하는 세 부분이 있다.

[0106] 도 2는 토마토 열매 맺는 시스템의 포즈 골격의 예를 보여준다. 여기서 오른쪽과 왼쪽은 각각 단일 및 다중 토마토의 골격이다. 관절을 공유할 수 없는 인간 포즈와 같은 인간 포즈 추정의 일반적인 골격과 달리 일부 토마토는 여러 토마토의 경우 연결된 꽃자루에서 가지 점을 공유할 수 있다.

[0108] 토마토 다발을 수확하는 것과 같은 다양한 수확 방법을 고려하여 더 많은 키포인트 유형을 정의하고 더 자세한 포즈 표현이 필요할 수 있다. 그러나 이것은 복잡성과 처리 시간을 증가시킬 수 있다.

[0110] 본 발명에서는 단일 과일 수확을 가정하여 탈리대(abcission zone)를 포함하여 4가지 유형의 키포인트로 기하학적 정보를 제공하는 것으로 충분하였다.

[0112] 2.2. 데이터세트 구성

[0114] 데이터 세트는 자체 수집한 이미지와 다운로드 가능한 공개 데이터 세트를 사용하여 구성되었다. 영상획득은 국내 토마토 재배를 위한 일반 수경온실에서 진행하였으며, 영상은 두 가지 방식으로 촬영하였다. 하나는 스마트폰(갤럭시S21, 삼성, 한국)을 수동으로 사용하여 촬영하는 것이고 다른 하나는 온실 내 온수 파이프를 원격으로 주행할 수 있는 개발된 모니터링 시스템으로 자동으로 사용하는 것이다.

[0116] 토마토 열매 맺는 시스템의 키포인트(이하, 토마토 포즈)를 포함하기 위해 다양한 시점에서 수동 캡처를 수행했으며 이미지의 해상도는 4000 × 3000 픽셀로 촬영되었다.

[0118] 수확 로봇의 관점에서 토마토 이미지를 연속적으로 포착하기 위해 모니터링 시스템 측면에 RealSense 카메라(D435, Intel, CA, USA)를 설치하고 시스템이 파이프의 경로를 따라 이동하도록 제어했다.

[0120] D435는 1280 × 720 픽셀의 해상도로 30fps로 이미지를 캡처했다. 그러나 캡처된 이미지는 연속 캡처된 이미지 사이에서 동일한 개체를 포함하는 유사한 장면을 제거하기 위해 3fps로 저장되었다. 총 녹화 시간은 온실에서 10개의 서로 다른 배드 열에 대해 100초였으며 이미지는 샘플링되었다.

[0122] 특히, 개방형 데이터셋인 작업 토마토(Laboro Tomato)는 다양한 관점에서 캡처한 다양한 숙성 단계의 토마토 이미지로 구성되며 객체 감지 및 인스턴스 분할 작업을 위한 주석이 제공된다. 이미지는 3024 × 4032와 175 3120 × 4160의 2가지 해상도를 가지고 있으며, 총 이미지 수는 9777개, 총 804개였다. 본 발명에서는 방울토마토를 제외한 443개의 토마토 이미지만을 사용하였다.

[0124] 포즈 추정을 위한 모델을 효율적으로 훈련시키기 위해 조도, 거리, 시점으로 인한 폐색으로 키포인트를 시각적으로 확인하기 어렵도록 이미지를 필터링했다. 물체를 중심으로 일부 영역을 잘라내어 이미지를 여러 샘플로 나

눴다.

[0126] 따라서 447개의 이미지를 샘플로 사용하여 모델 학습용 데이터셋을 구성하고 자체 개발한 주석 도구를 사용하여 각 이미지별로 키포인트를 2D 좌표(x, y)로 표시했다.

[0128] 레이블이 지정된 샘플을 7:3 비율로 훈련 세트와 검증 세트로 무작위로 나누고 검증 세트의 절반을 테스트 세트로 사용했다. 샘플 수는 훈련, 검증 및 테스트 세트에 대해 각각 337, 68 및 68이었습니다. 또한 모델을 훈련할 때 각 키포인트의 위치에 가우시안 노이즈를 추가하는 데이터 증대 기법을 사용했다.

[0130] 2.3. 모델 및 교육

[0131] OpenPose는 토마토 열매 맺는 시스템을 위한 자세 추정 프레임워크를 구성하기 위한 백본으로 사용되었다. OpenPose는 키포인트 간의 연관에 대한 비모수적 표현인 PAF(부분 친화성 필드)를 활용하는 다중인용 2D 포즈 추정 방법이다. 이전 연구에서 저자는 PAF를 기반으로 키포인트의 탐지 성능을 개선하고 글로벌 컨텍스트를 반영하여 여러 사람 포즈를 구문 분석했다. 손과 발과 같은 사람의 키포인트는 다른 키포인트와의 연결에 관계없이 이미지와 구별될 수 있다.

[0133] 한편, 연결 없이 토마토 가지의 키포인트를 추정하는 것은 어렵다. 따라서 PAF와 같은 키포인트의 연관성을 추정하는 것이 필수적이다. 이 하위 섹션에서는 OpenPose를 기반으로 하는 토마토의 포즈 추정 방법에 대해 간략하게 설명한다.

[0135] 2.3.1. 신뢰맵(Confidence Maps) 및 부분 선호도 필드(Part Affinity Fields)

[0137] 토마토의 포즈 추정에는 두 가지 유형의 2D 표현이 사용된다. 이는 "Z. Cao, T. Simon, S.-E. Wei, Y. Sheikh, Realtime multi-person 2d pose estimation using part affinity fields, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 7291-7299."에서 제안한 대로 신뢰 맵과 PAF이다.

[0139] 첫째, 신뢰 맵은 키포인트의 분포를 히트 맵으로 나타낸다. 이미지에 있는 K 토마토의 경우 각 토마토에는 J 키포인트가 있다.

[0140] 이미지의 신뢰 맵은 각 키포인트의 유형에 해당하는 J개의 채널로 구성되며, 하나의 이미지에 포함된 모든 토마토에 대한 특정 키포인트의 위치는 신뢰 맵의 각 채널에 표시된다.

[0142] 신뢰 맵의 각 픽셀 좌표 p ∈ R2에 대해 식 1과 같이 정답 S*를 생성한다.

$$\mathbf{S}^*_{j,k}(p) = \exp\left(-\frac{\|\mathbf{p}-\mathbf{x}_{j,k}\|_2^2}{\sigma^2}\right) \qquad (1)$$

[0143]

[0144] 여기서 j ∈(1, 2, ..., J)는 키포인트의 종류에 따른 표기법이고 k ∈(1, 2, ...,K)는 영상에서 토마토의 숫자 표기법이다. σ는 각 키포인트 위치의 신뢰도를 확산하는 데 사용되는 하이퍼파라미터이다.

[0146] 각 이미지의 모든 토마토에 대한 키포인트 j에 대한 집계된 신뢰 맵은 최대 연산을 통해 식 2로 표시된다.

$$\mathbf{S}^*_j(p) = \max_k \mathbf{S}^*_{j,k}(\mathbf{p}) \qquad (2)$$

[0147]

[0148] 이러한 신뢰 맵은 섹션 2.3.2에 설명된 신경망을 사용하여 훈련되고 추정되었다. 각 토마토의 포즈는 PAF를 사용하여 감지된 일련의 키포인트에서 추정할 수 있다.

[0150] 이것은 토마토의 부분(즉, 키포인트의 연결)을 2차원 벡터 필드로 나타내고 부분의 위치와 방향에 대한 정보를 포함하며 각 토마토에 대해 C개의 연결을 가진다.

[0152] 두 키포인트 j1, j2의 위치 xj1,k, xj2,k를 연결하는 부분 c의 PAF L*c,k는 단위 벡터 v를 이용하여 수학식 3과 같이 정의된다.

$$\mathbf{L}^*_{c,k}(\mathbf{p}) = \begin{cases} \mathbf{v} & \text{if } \mathbf{p} \text{ on limb } c,k \\ \mathbf{0} & \text{otherwise} \end{cases} \quad (3)$$

[0154] $$\mathbf{v} = (\mathbf{x}_{j_2,k} - \mathbf{x}_{j_1,k}) / \|\mathbf{x}_{j_2,k} - \mathbf{x}_{j_1,k}\|_2$$

[0156] 여기서 각 부분의 점집합 p는 식 4와 같이 두 키포인트를 연결하는 선분으로부터 수직거리 σl보다 작거나 같은 점으로 정의된다.

[0157] $$0 \le \mathbf{v} \cdot (\mathbf{p} - \mathbf{x}_{j_1,k}) \le l_{c,k} \text{ and } |\mathbf{v}_\perp \cdot (\mathbf{p} - \mathbf{x}_{j_1,k})| \le \sigma_l \quad (4)$$

[0158] 여기서 $l_{c,k} = \|\mathbf{x}_{j_2,k} - \mathbf{x}_{j_1,k}\|_2$ 는 부분의 길이이고 v⊥는 v에 수직인 벡터이다.

[0160] 실제 PAF는 식 5에 표시된 대로 각 이미지에 있는 모든 토마토의 부분 c에 대한 PAF의 평균이다.

[0161] $$\mathbf{L}^*_c = \frac{1}{n_c(\mathbf{p})} \sum_k \mathbf{L}^*_{c,k}(\mathbf{p}) \quad (5)$$

[0163] 여기서 nc(p)는 이미지에 있는 모든 토마토의 c 부분에 대한 점 p에서 0이 아닌 PAF 값의 수이다.

[0165] 2.3.2. 네트워크 아키텍처

[0167] 본 발명에서는 토마토 포즈 추정을 위해 도 3에 묘사된 경량 OpenPose 아키텍처를 사용했다. 특징 F 추출을 위한 백본, 신뢰 맵 및 PAF의 초기 추정을 위한 초기 단계, 보다 정확한 추정을 위한 T-1 정제 단계로 구성된다. 각 단계에 대해 신뢰 맵의 집합 S = (S1, S2, ..., SJ)와 PAF의 집합 L = (L1,L2, ...,LC)가 예측되었다. 각 단계의 예측 결과는 특징 맵 F와 연결되어 다음 단계의 입력으로 사용되어 보다 정교한 결과를 예측할 수 있다.

[0169] 그라운드트루(groundtruth) 신뢰 맵 S^*_j 및 PAF L^*_c 를 사용하여 식 6과 같이 네트워크 훈련을 위한 손실 함수를 적용했다.

[0170] $$f^t_{\mathbf{S}} = \sum_{j=1}^{J} \sum_{\mathbf{p}} \mathbf{W}(\mathbf{p}) \cdot \|\mathbf{S}^t_j(\mathbf{p}) - \mathbf{S}^*_j(\mathbf{p})\|_2^2$$
$$f^t_{\mathbf{L}} = \sum_{c=1}^{C} \sum_{\mathbf{p}} \mathbf{W}(\mathbf{p}) \cdot \|\mathbf{L}^t_c(\mathbf{p}) - \mathbf{L}^*_c(\mathbf{p})\|_2^2 \quad (6)$$
$$f = \sum_{t=1}^{T} f^t_{\mathbf{S}} + f^t_{\mathbf{L}}$$

[0172] 여기서 t는 초기 단계, Stj(p), Lt를 포함한 단계 번호이고, $\mathbf{S}^t_j(\mathbf{p}),\ \mathbf{L}^t_c(\mathbf{p})$는 신뢰 맵이고 PAF는 각 단계에서 추정된다. W(p)는 주석이 누락된 픽셀 p에 의해 참양성(true positive)에 불이익을 주는 것을 피하기 위해 사용되는 이진 마스크이다. 멀티 토마토 포즈 추정은 학습된 모델에 의해 예측된 최종 단계의 신뢰 맵 S와 PAF L을 사용하여 수행되었으며, 이에 대해서는 다음 하위 섹션에서 설명한다.

[0174] 2.3.3. PAF 기반 멀티 토마토 포즈 추정

[0176] 토마토 포즈 추정은 신뢰 맵에서 후보 키포인트 위치를 획득하고 PAF를 기반으로 이들 간의 연결을 찾는 방식으로 수행된다.

[0178] 먼저, 예측된 신뢰 맵에서 피크 포인트를 탐지하고 NMS(Non-Maximum Suppression)를 수행하여 노이즈를 제거하여 후보 키포인트 위치를 추정했다. 또한 여러 토마토의 키포인트가 하나의 이미지에 함께 있기 때문에 PAF를 사용하여 각 토마토의 연결을 포함하여 전체 포즈를 추정한다.

[0180] 부분에 포함된 두 개의 키포인트 후보의 연관 점수는 식 7과 같이 두 후보 위치 dj1, dj2를 연결하는 선분을 따라 해당 PAF의 적분을 계산하여 구한다.

$$E = \int_{u=0}^{u=1} \mathbf{L}_c(\mathbf{p}(u)) \cdot \frac{\mathbf{d}_{j_2} - \mathbf{d}_{j_1}}{\|\mathbf{d}_{j_2} - \mathbf{d}_{j_1}\|_2} du \qquad (7)$$

$$\mathbf{p}(u) = (1-u)\mathbf{d}_{j_1} + u\mathbf{d}_{j_2}$$

[0182]

[0183] 여기서 p(u)는 두 개의 키 포인트 dj1 및 dj2의 보간이다.

[0185] 각 부분 유형에 대한 연결 가능한 모든 키포인트 후보에 대해 연관 점수를 얻었고 NMS를 수행했다. 구체적으로, 연관 점수가 높은 순서대로 부분을 연결하고, 한번 연결이 확인된 키포인트 후보는 다른 키포인트와의 연결 후보에서 제외한다.

[0187] 그러나 인간과 달리 토마토는 줄기 키포인트를 공유한다. 꽃자루는 하나 이상의 소화경(작은 꽃자루)을 가질 수 있다. 따라서 근위 소화경 부분 추정에 NMS를 적용하는 것은 적합하지 않다.

[0189] 따라서 예외적으로 토마토 포즈의 마지막 부분인 근위 소화경은 NMS를 사용하지 않고 가장 높은 연관 점수를 가진 가지 점(키포인트 4)을 연결하여 여러 토마토의 부분이 공유할 수 있도록 한다.

[0191] 마지막으로 4개의 키포인트와 토마토의 3개 부분으로 완전한 포즈의 추정 결과를 활용했다.

[0193] 2.4. 평가

[0195] 본 발명의 접근 방식의 성능을 평가하기 위해 감지된 키포인트와 부분을 사람이 미리 결정한 실제와 비교했다. 검출된 점과 2차원 영상면의 실측면 사이의 픽셀 차이인 키포인트 오차는 픽셀 수준에서 계산되었다.

[0197] 두 개의 연속적인 키포인트인 원위 소화경과 근위 소화경 사이의 선형 연결은 실제 길이와 각도 차이를 비교하여 평가되었다. 여기서 길이 오차는 픽셀 수준에서의 절대 차이이고 방향 오차는 두 선에 의해 형성된 예각이다.

[0199] 또한 식 8에 정의된 PDK(검출된 키포인트)의 백분율을 측정했다. PDK는 추정된 키포인트가 ground truth로부터 임계값 msi 내에 있는 경우 true positive로 간주하여 탐지 정확도를 평가한다.

$$PDK_j = \frac{\Sigma_{i=1}^{N_j} \delta(d_{ij} \geq ms_i)}{N_j} \qquad (8)$$

[0200]

[0201] 여기서 δ는 조건이 참이면 1을 반환하고 거짓이면 0을 반환하는 임계값 함수이다. dij는 i번째 객체에서 j번째 키포인트의 예측된 위치와 실제 위치 사이의 거리이다. si는 토마토 과일-베어링 시스템의 경계 상자의 대각선인 객체 i의 척도이다. m은 객체 si의 스케일에 대한 정규화된 거리이고, Nj는 j번째 키포인트의 총 수이다.

[0203] 3. 결과 및 논의

[0204] 3.1. 멀티 토마토 포즈 추정

[0206] 도 4는 멀티 토마토 포즈 추정의 대표적인 결과를 보여준다. 각 토마토의 추정 포즈는 열매와 소화경의 자세를 나타내며 4개의 키포인트가 순차적으로 연결된 골격으로 시각화된다. 오브젝트에 포함된 4개의 키포인트는 서로 구별하기 위해 서로 다른 색상으로 표현된다.

[0208] 과실 중심, 꽃받침, 탈락대, 가지 점은 각각 청색, 녹색, 적색, 황색으로 플로팅하였으며, 연속되는 2개의 키포인트인 부분의 색상은 시작 키포인트과 동일하였다. 예를 들어, 말단 소화경은 꽃받침과 동일한 녹색을 띠고 있다(키포인트 2). 첫 번째 행은 단일 오브젝트(토마토)의 예이고, 아래쪽 행은 여러 토마토에 대한 포즈 추정 결과이다. 행이 낮을수록 한 번에 감지할 수 있는 개체의 수가 커진다. 대부분의 경우 검출된 토마토 포즈는 토마토 개수에 상관없이 합리적으로 추정되어 만족스러운 성능을 보였다. 그러나 추정된 위치는 실제와 유사하지만 오류가 있는 키포인트를 감지하는 등의 가양성(false positive)이 있었다.

[0210] 이러한 문제의 대부분은 토마토가 작게 나타나거나 뭉쳐져 있어 인간이 직관적으로 판단하기 어려운 이미지에서 발생한다. 결과는 부정확한 경우를 포함하지만 수확 로봇이 움직임을 통해 이러한 조건보다 가까운 물체를 인식한다는 점에서 수확에 대한 포괄적인 정보를 제공할 수 있다.

[0212] 도 5는 토마토 개체 간의 가지 점 공유에 대한 근위 소화경부를 추정하기 위해 NMS 방법을 적용하지 않은 효과의 결과를 나타낸다. 각 행의 2개의 이미지는 동일한 장면에 대한 이미지이고, 오른쪽 열의 도은 가지 점을 공

유하여 추정한 포즈이고, 왼쪽 열의 도은 NMS 방법을 사용하여 얻은 결과이다.

[0214] 본 발명의 접근 방식은 꽃자루의 가지 지점에서 작은 꽃자루가 공유되는 토마토 구조의 형태를 정확하게 나타낼 수 있다.

[0216] 3.2. 예상 토마토 포즈 분석

[0218] 또한 제안된 접근 방식의 키포인트 오류와 부분 오류를 정량적으로 평가했다. 키포인트 오류는 표 1에 나열되어 있으며 오류는 키포인트 유형(과중심, 꽃받침, 탈락대, 가지 점)으로 표시된다.

[0220] [표 1] 감지된 키포인트 및 PDK@0.5 점수의 위치 오류.

[0221]

Axis	Fruit-center ↓	Calyx ↓	Abscission zone ↓	Branch-point ↑	Total ↑
x (pixel)	7.8±5.18	4.2±3.10	6.9±5.91	13.0±17.91	8.0±8.03
y (pixel)	9.6±5.64	3.8±2.72	7.8±5.38	17.6±25.34	9.7±9.77
PDK@0.5 (%)	92.6	99.9	99.4	79.4	92.9

[0223] 오차는 픽셀 단위로 계산하였으며, 그 범위는 가로로 7.8-13.0, 세로로 3.8-17.6으로 관찰되었다. 모든 키포인트의 평균 오차는 x축과 y축에 대해 각각 8.0과 9.7이었다. 이러한 오차는 D435 카메라를 사용할 때 0.5m 거리에 있는 물체에 대해 수평으로 4.3mm, 수직으로 5.2mm에 불과하다. 픽셀 오차는 약 4.0으로 꽃받침에서 가장 작았고, 가지 점은 13.0-17.6의 오차 범위로 가장 낮은 성능을 보였다. 가지 점의 키포인트 오차는 다른 키포인트보다 3~4배 높았다. 이는 가지가 다른 지역의 줄기 부분과 구별하기 어려운 것으로 해석된다. 그러나 꽃받침과 탈리대의 픽셀 오차는 가장 낮았다. 이 두 가지 핵심 포인트는 로봇이 수확을 위한 피킹 위치를 결정하는 데 기하학적으로 중요하다.

[0225] 표 1의 마지막 행은 0.5의 정규화 거리에서의 PDK를 보여준다. 여기서 검출된 키포인트는 오차가 경계 상자의 대각선의 절반 미만일 때 정확하고 오차는 검출된 위치와 그라운드트루(ground truth) 사이의 유클리드 거리이다. 모든 키포인트에 대한 평균 PDK는 92.9%였으며, 꽃받침과 이탈부가 가장 높은 검출 성능과 실용적인 사용성을 나타냈다.

[0227] 표 2는 과실축, 원위 소화경 및 근위 소화경에 대한 길이 및 각도 오차로 표현되는 부분 오차를 나타낸다.

[0229] 표 2: 추정 쌍의 길이 및 각도 오차 및 PDK@0.5 점수.

[0230]

Item	Fruit-axis ↓	distal pedicel ↓	Proximal pedicel ↓	Total ↑
Length (pixel)	8.0±3.95	7.4±4.83	15.2±16.13	10.2±8.30
Angular (deg)	21.3±21.86	20.6±17.27	18.0±17.56	20.1±18.90
PDK@0.5 (%)	92.6	99.4	78.9	90.3

[0232] 길이 오차는 실측값과 예측값 사이의 픽셀 단위 길이 차이로 오차 범위는 7.4~15.2, 평균은 10.2였다. 이 평균 길이 오차는 실제 사용에 적합한 D435 카메라로 0.5m 거리에서 피사체를 촬영할 때 실제 길이로 약 5.5mm에 불과하다. 각도 오차는 실제 부분과 예측 부분 사이의 예각이며 아크탄젠트를 사용하여 계산되었다. 각도 오차는 18.0° - 21.3° 범위였으며 평균값은 약 20° 였다. 가장 작은 오차는 길이 차이가 7.4인 원위 소화경과 18.0°의 각도 차이를 가진 근위 소화경에서 관찰되었다.

[0234] 부분 추정의 성능도 PDK를 사용하여 평가되었다. 전반적으로 해당 키포인트의 탐지율이 높은 부분이 더 좋은 성능을 보였다. 예를 들어, 원위 소화경은 99%의 가장 높은 검출 성능을 보인 반면, 가지 점을 포함하는 근위 소화경은 80% 미만의 점수를 나타냈다. 이는 두 소화경 모두 탈리대인 높은 검출율 키포인트를 포함하는 반면, 근위 소화경은 낮은 검출률을 갖는 가지 점을 포함하기 때문이다. 과실축 부분에서도 90% 이상의 검출 성능을 나타냈다.

[0236] PDK는 정규화된 거리 k에 따라 변한다. 따라서 PDK는 정규화 거리로 표현되는 토마토 과실 맺는 시스템의 대각선에 대한 비율인 임계값 수준에 따라 평가되었으며 0-1.0 범위에서 0.1 증분으로 설정되었다.

[0238] 도 6은 정규화된 거리에 따른 검출 정확도를 보여준다. 키포인트에서 꽃받침(키포인트 2)은 PDK@0.2에서 약 95%의 높은 성능을 보였고, 탈락대는 약 81%로 다음으로 가장 높았다. 가지 점을 제외한 모든 키포인트는 0.4의 정규화 거리에서 90%에 가깝거나 이상으로 관찰되었으며, 특히 꽃받침과 탈리대는 거의 완벽한 수준을 보였다.

[0240] 가지 점은 0.2 이상의 정규화 거리에서 가장 낮은 점수를 보였으며 1.0에서 최대 94%에 도달했으며 다른 키포인트는 0.5의 정규화 거리에서 이미 해당 값에 도달했다.

[0242] 부분의 점수는 원위 소화경이 가장 높은 검출률을 보였으며 PDK@0.3이 약 95%인 반면 다른 부분은 80% 미만의 점수를 보였다. 과일 부분도 0.4 이상의 임계값 수준에서 90% 이상의 탐지 성능을 보였다. 근위 소화경의 경우 가지 점에서의 낮은 검출 성능으로 인해 임계값 0.4에서 70%에 가까웠다.

[0244] 길이와 각도 오차로 표현되는 추정된 포즈의 성능은 도 7에서 히스토그램으로 표현된다. 검출된 부분을 오차로 카운트하고, 각 카운트를 전체 개체수로 나누어 확률을 표현하였다. 왼쪽과 오른쪽 기둥은 각각 길이와 각도 오차이며, 기둥은 위에서부터 토마토 열매, 원위 소화경, 근위 소화경의 순서로 표현하였다. 각 그래프는 갈색과 녹색 막대로 각각 표시되는 확률 분포와 누적 분포로 구성된다. 수평 화살표 라인은 높은 정확도로 모든 샘플에 대해 50%의 최대 오류를 나타낸다.

[0246] 결과는 토마토 열매, 원위 소화경, 근위 소화경에 대해 길이 7, 6, 8 픽셀에서 최대 오차가 관찰되었고 각도 10, 14 및 9° 에서 관찰되었다. 평균적으로 추정된 쌍의 절반은 7픽셀 내에서 충족되고 길이 및 각도 오류는 11° 이다.

[0248] 3.3. 토론

[0249] 본 발명의 방법은 이미지에 작은 꽃자루가 있는 여러 토마토에 대한 포즈 추정을 동시에 제공할 수 있다. 또한 기존 연구에서는 다단계 프레임워크로 구성되어 다중 객체 감지를 위해 하향식 방법을 사용했지만 본 발명은 종단 간 프로세스가 포함된 상향식 방법을 사용했다.

[0251] 사용된 방법은 실용적이고 효율적인 토마토 포즈 추정을 제공할 뿐만 아니라 동시에 추정할 토마토의 수가 증가하더라도 비례적으로 계산량이 증가하지 않았다.

[0253] 본 발명의 방법의 성능은 검출률이 약 0.93으로 나타났으며, 점수가 0.93-0.88 범위인 최근의 인간 포즈 추정뿐만 아니라 최근의 동물 포즈 추정에서도 유사한 결과를 보였다. 점수는 소와 닭에 대해 각각 0.90-0.94 및 0.63-0.96이다.

[0255] 인간 포즈 추정의 키포인트 구성은 좀 더 복잡하지만, 본 발명의 결과는 기초 자료가 부족하고 사물 간의 구분이 어려워 처음 토마토 포즈 추정에 적용한다는 점에서 의의가 있다.

[0257] D435 카메라를 사용할 때 0.5m 거리에 있는 물체에 대해 평균 5.2mm 미만의 오차로 키포인트가 감지되었다. 따라서 시각적 서보와 같은 시각적 기반 제어 장치에 사용할 수 있다.

[0259] 검출된 특징점의 백분율(PDK)을 이용하여 검출 성능을 평가하였으며, 과실 중심, 꽃받침, 탈락대, 가지 점의 특징점은 PDK@0.5가 각각 0.93, 0.99, 0.99, 0.74였다. 부분 검출의 경우 PDK@0.5는 과실축, 원위 소화경 및 근위 소화경에 대해 각각 0.93, 0.99 및 0.79였다. 가지 점을 제외한 점수는 0.94-0.97 범위로 토마토 다발 탐지와 유사한 범위를 가지며 0.91로 꽃자루 위치 탐지보다 높은 성능을 보여주었다.

[0261] 특히, 절단점을 결정하는 중요한 검출 단서이자 과일 박리와 관련된 탈리대 및 원위 소화경에서 높은 검출률을 나타내어 따기 로봇에 적용한다는 점에서 의미 있는 결과를 보였다.

[0263] 검출된 토마토 부위의 자세에서 길이와 각도오차는 각각 평균 8.3픽셀, 18.9° , 중앙값 7픽셀, 11° 로 관찰되었다. 결과는 23.4° 의 각도 오차로 종래의 자세 추정보다 우수한 성능을 나타내었지만 5° 미만의 각도 오차를 갖는 꽃자루 탐지보다 낮은 성능을 나타냈다.

[0265] 4. 결론

[0266] 본 발명은 토마토 과일 줄기 시스템의 포즈를 추정하고 부분 친화성 필드를 기반으로 여러 토마토 과일 소작경쌍을 동시에 감지했다. 이 방법의 장점은 계산 부하가 감지할 개체의 수를 늘리는 데 크게 의존하지 않고 프레임워크가 엔드 투 엔드 프로세스로 구성되어 다중 포즈 토마토를 추정할 수 있다는 것이다. 특히, 수확의 절단점으로 중요하고 자주 사용되는 꽃받침, 절단부, 원위 소화경이 잘 추정되었다.

[0268] 본 발명의 접근 방식은 절단 유형 및 따기 유형 로봇 수확기 모두에 대해 수확 가능한 과일에 대한 유용한 정보

를 제공할 수 있다. 본 발명에서 제공한 여러 토마토 열매 소화경의 자세 정보는 절단 위치 선택과 같은 수확 메커니즘뿐만 아니라 로봇 수확 중 접근성 평가를 기반으로 수확 우선 순위를 결정하는 데에도 활용할 수 있다.

[0270] 도 8은 전술한 바와 같은 과일 줄기 자세를 인식하는 과정을 포함하는 토마토 포즈 추정 시스템의 일례를 나타낸다. 도 8에 도시된 바와 같이 토마토 포즈 추정 시스템의 경우 대상이되는 과일, 예를 들면 토마토의 영상을 촬영하는 영상 취득부(100); 취득 영상으로부터 토마토의 골격 정보(예를 들면 과일-줄기의 포즈 정보)를 추정하는 과일-줄기 포즈 추정부(300) 및 추정된 과일-줄기 정보를 수확 로봇 등으로 전송하기 위한 추정 정보 전송부(400)를 포함하고, 과일-줄기 추정부(300)는 인공지능 학습모델(200)을 통해 전술한 바와 같은 과일-줄기 추정 동작을 수행하도록 학습되어 있다.

[0272] 영상 취득부(100)는 영상센서와 제어기로 구성되며 영상센서는 RGB일 경우 2D 자세 정보의 제공이 가능하며, depth가 포함될 경우 3D 정보로 확장이 가능하도록 구성된다.

[0274] 과일 줄기 추정부에서, 추정되는 과일-줄기 쌍의 주요부위는 토마토의 경우 과일의 중심부(fruit center), 과일과 줄기가 연결되는 꽃받침(calyx), 소화경(pedicel) 중심에 위치한 탈리대(abscission zone), 소화경과 화방(peduncle)의 연결점(branch point)으로 선정된다(이를 주요부위(키포인트) 1, 2, 3, 4라 함).

[0276] 주요부위 검출은 학습 모델(200)을 통해 지도학습 기반의 알고리즘이 사용되며 주요부위가 라벨링 된 다양한 토마토-줄기 영상들을 이용하여 학습된다. 이때 영상 취득부(100)를 통해 촬영된 영상을 통해 학습이 이루어질 수도 있다.

[0278] 학습시 과일-줄기쌍은 하나의 영상 내 복수개가 존재하므로 PAF (part affininty field) 등의 주요부위 연결강도 학습방법을 이용하여 개체별 1-2-3-4 연결강도를 학습시킨다.

[0280] 학습 모델을 이용하여 입력된 영상에 존재하는 다양한 과일-줄기쌍의 모든 주요부위들의 위치를 우선적으로 검출한 후 연결강도를 이용하여 개체별로 1-2-3-4 순으로 주요부위를 선형적으로 연결하며, 이렇게 연결된 1-2-3-4의 골격(skeleton)을 토마토-줄기의 자세로 평가하게 된다.

[0282] 각각의 자세의 기하학적 위치는 과일-줄기 추정부(300)에서 4개의 검출된 주요부위의 2D pixel 위치와 관련되어 있으며, depth 정보나 단편적인 거리 측정을 통해 3D 실제좌표계로 변환이 가능하다.

[0284] 이렇게 추정된 2D 또는 3D의 기하학적 정보는 추정정보 전송부(400)를 통해 수확 로봇으로 전송되고, 수확 로봇은 이를 이용하여 수확로봇의 그리퍼가 잡을 수 있는 위치, 자세의 과일-줄기의 선정이 가능하여, 이를 기반으로 수확 메커니즘 전략을 수립할 수 있게 된다.

[0286] 이상에서 설명된 장치는 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성요소는, 예를 들어, 프로세서, 콘트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPGA(field programmable gate array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및하나의 콘트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.

[0288] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively)처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상 장치(virtual equipment), 컴퓨터 저장매체 또는 장치에 구체화(embody)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

[0290] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판

독 가능 매체에 기록될 수 있다. 상기 컴퓨터판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다.

[0292] 이상과 같이 실시예들이 비록 한정된 실시예와 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기의 기재로부터 다양한 수정 및 변형이 가능하다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.

[0294] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속하는 것으로 해석되어야만 한다.

부호의 설명

[0296] 100: 영상 취득부

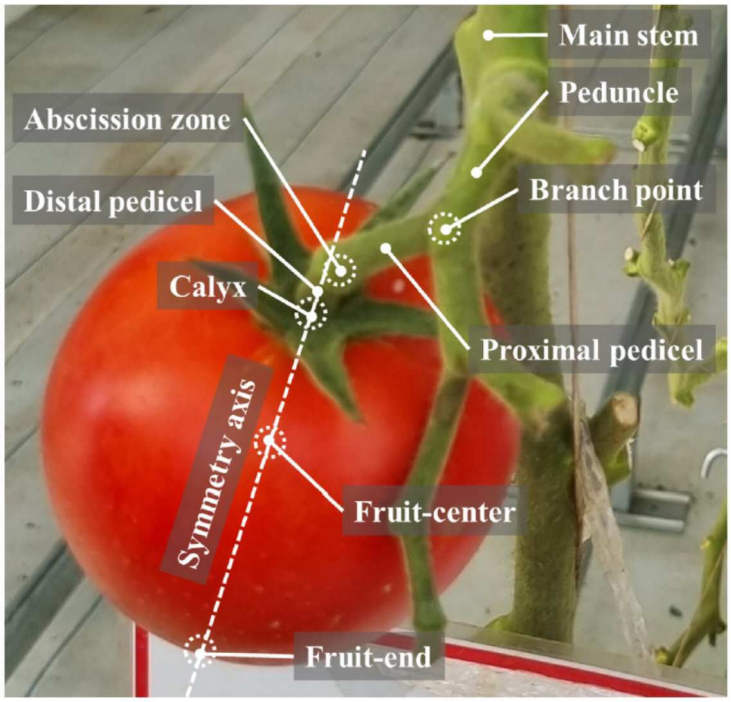
200: 학습 모델

300: 과일-줄기 추정부

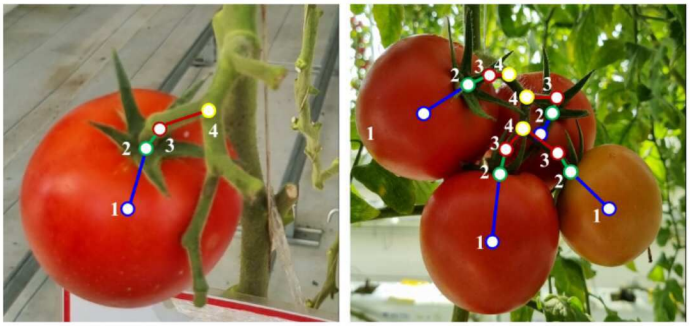
400: 추정정보 전송부

도면

도면1

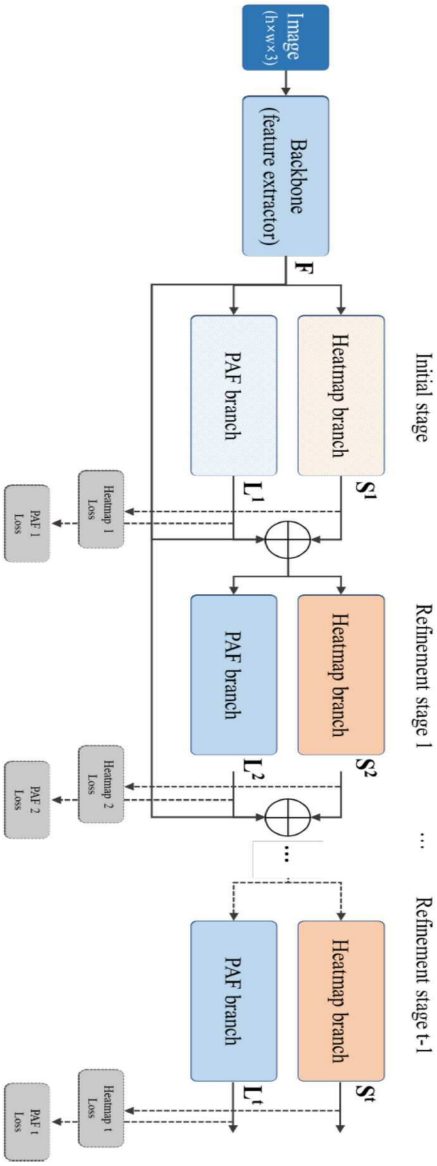


도면2



(1) Fruit-center, (2) Calyx, (3) Abscission zone, (4) Branch point

도면3



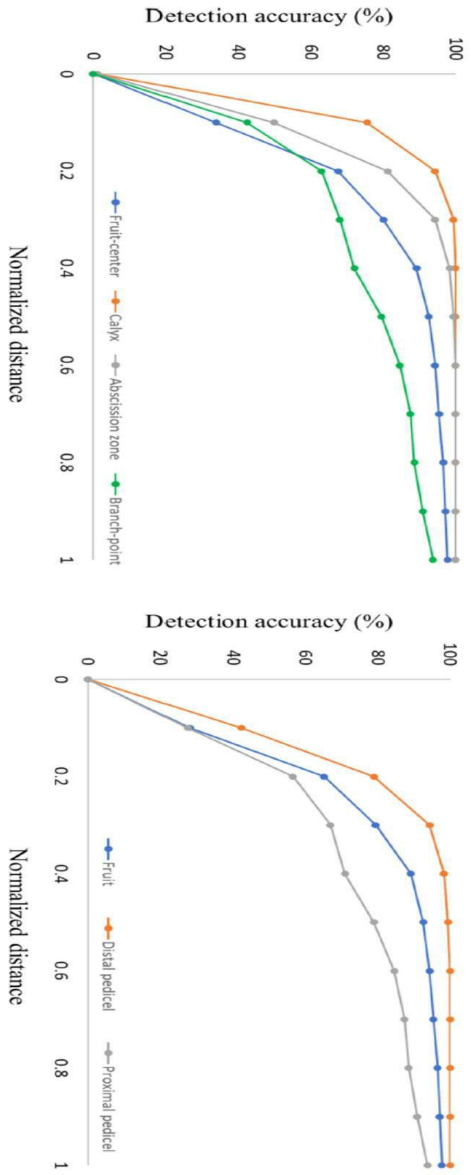
도면4



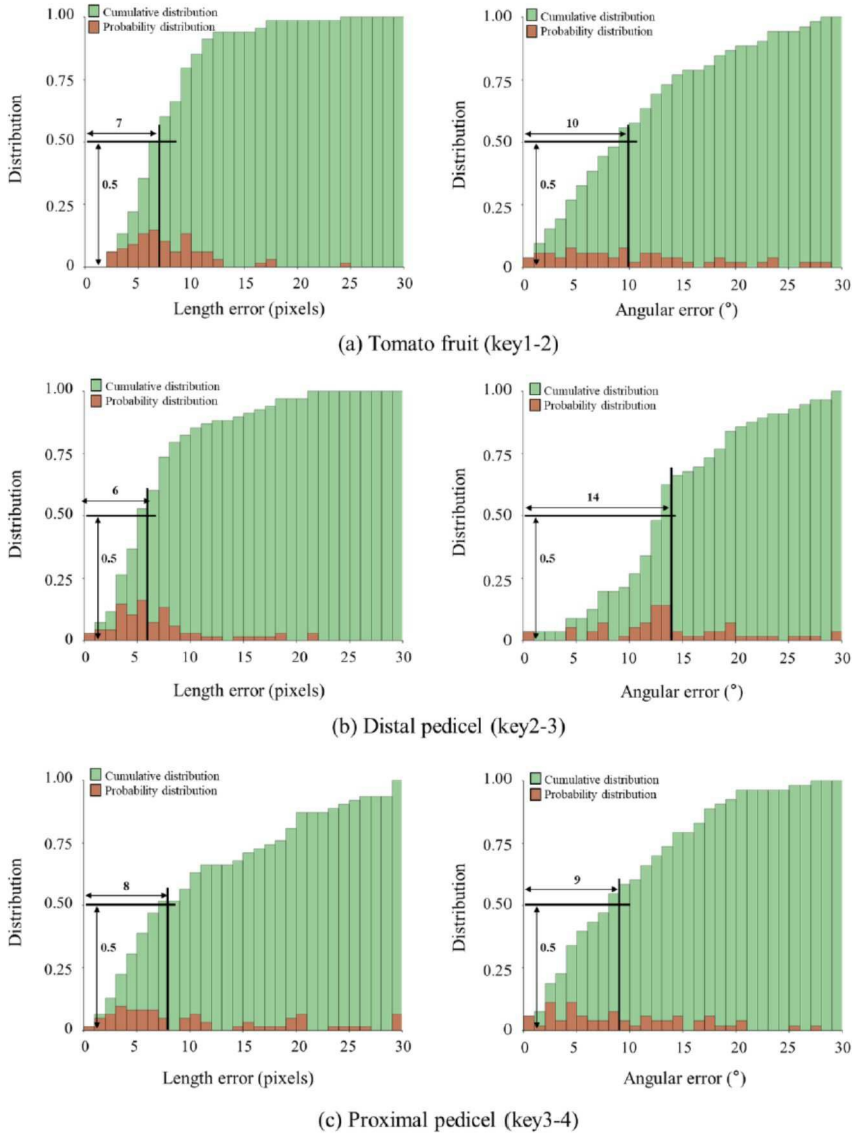
도면5



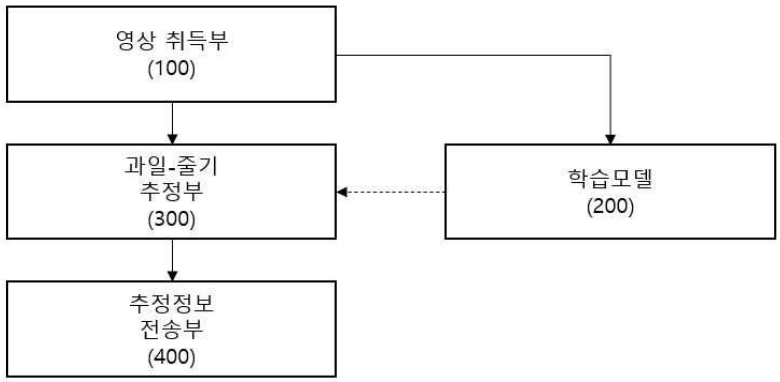
도면6


Detection accuracy (%)
0
20
40
60
80
100
Normalized distance
0
0.2
0.4
0.6
0.8
1
Fruit-center
Calyx
Abscission zone
Branch-point
Detection accuracy (%)
Normalized distance
Fruit
Distal pedicel
Proximal pedicel

도면7


Cumulative distribution
Probability distribution
Distribution
Length error (pixels)
Angular error (°)
7
10
6
14
8
9
0.5
(a) Tomato fruit (key1-2)
(b) Distal pedicel (key2-3)
(c) Proximal pedicel (key3-4)

도면8


영상 취득부
(100)
과일-줄기
추정부
(300)
학습모델
(200)
추정정보
전송부
(400)