

Vision-Based Tracking of Construction Entities

-2013 글로벌 CEM R&D Trend 세미나 발표 내용-



박만우 명지대학교 토목환경공학과 조교수

1. 연구 배경

Tracking 시스템은 추적 개체들의 위치를 시간이력으로 제공한다. 일반적 건설 현장의 복잡성, 밀집도, 그리고 다양한 개체들의 움직임을 고려해 볼 때, tracking 시스템이 제공하는 데이터는 효율적 건설 관리 및 현장 모니터링에 중요한 정보로 활용될 수 있다. 건설 인부들과 여러 종류의 건설 장비들, 그리고 건설 자재들의 위치를 실시간으로 구할 수 있다면, 진도 관리, 현장 안전 관리, 생산성 분석, 공정 관리 등의 신뢰도를 높이고 체계적 관리를 이루는 데에 이바지할 수 있을 것이다. 건설현장에서의 tracking 시스템으로서 무선 주파수 기술을 기반으로 한 RFID (Radio Frequency Identification), GPS (Global Positioning System), UWB (Ultra Wide-Band) 등이 주목을 받아왔고, 꾸준한 연구가 진행되어 왔다. 다양한 실험과 현장 테스트를 통해 이러한 시스템들의 적용성과 활용도가 발표되었고, 실제 건설 프로젝트에 적용된 사례들도 보고되었다. GPS는 건설 중장비들의 추적을 위해 사용되었고, RFID는 프리캐스트 콘크리트 부재들을 공장제작에서부터 현장 거치까지 추적하는 데에 사용된 바가 있다. 하지만, 무선 주파수 기반의 시스템에서는 추적하고자 하는 개체마다 센서(sensor)나 태그(tag)를 부착하여야 한다는 번거로움이 있다. 추적 개체 수가 늘어남에 따라 소요비용이 늘어나게 되며, 수시로 센서들을 회수하고, 다시 부착하는 번거로움을 감수해야 한다. 특히, 현장 인부들을 추적하기 위하여 센서를 부착하려 할 경우, 인부들로부터 강한 거부감과 반발을 야기할 수 있으며, 인부들이 센서를 임의로 제거할 시에도 규정 상 달리 취할 방도가 없다. 대규모의 건설장비 및 인력이 투입되는 건설 현장일수록, 이

러한 단점은 보다 부각되게 된다. 따라서 수많은 장비와 인력 등으로 밀집된, 넓은 규모의 현장에서는 이러한 기술들의 효율성이 떨어지게 된다. 이러한 경우를 위해, 추적 가능한 개체의 수를 최대화 하고, 시간 및 비용 면에서 효율성이 높은 기술이 요구된다.

Vision-based tracking 시스템은 센서나 태그를 필요로 하지 않으며, 비디오를 찍기 위한 카메라와 비디오를 입력 받아 프로세싱을 수행할 프로세서만이 요구될 뿐이다. 축구 중계에서 선수들의 주요 활동 지역과 주행거리를 산정하거나, 테니스 중계에서 공의 아웃판정을 하는 데에 이러한 기술들이 적용되고 있다. 이러한 솔루션은 넓고 밀집된 현장에서의 tracking을 위한, 보다 효율적인 대안으로서 충분한 활용 가치가 있다. 축구장을 건설 현장으로, 축구 선수들을 건설 장비나 인력으로 대체해 봄으로써, vision-based tracking 시스템의 적용 가능성을 엿볼 수 있을 것이다 (그림 1).

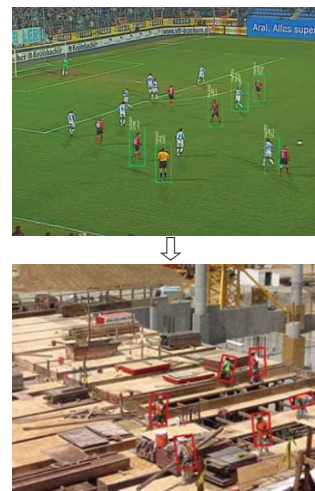


그림 1. Vision-based tracking의 적용성

컴퓨터 비전(computer vision) 분야의 기술들이 진화되고, 카메라 및 CPU의 성능이 지속적으로 발전하면서 vision-based tracking 시스템의 적용성과 가능성은 커지고 있다. 이러한 tracking 기술은 앞서 언급된 무선 주파수 기반 시스템들의 단점들을 보완할 수 있을 것으로 기대된다. 사용되는 장비를 고려해볼 때 비용 면에서 효율적이며, 장비의 설치 역시 간단하여 사용이 용이하고 시간 면에서도 효율적이다.

2. 시스템 구성 및 방법론

2.1 시스템 framework

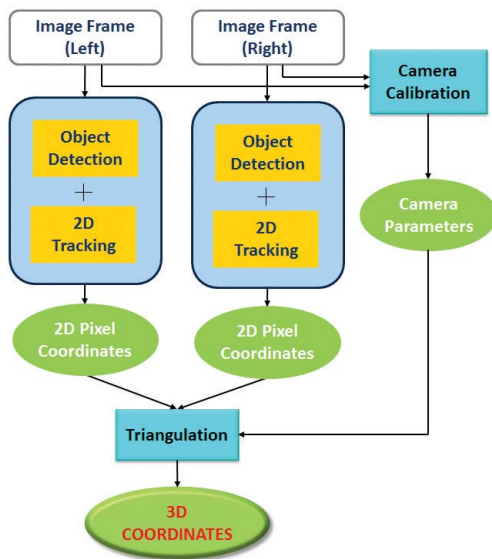


그림 2. 3D vision-based tracking framework

그림 2는 vision-based tracking 시스템의 전체 framework를 도식화한 것이다. 이 시스템은 기본적으로 2개의 고정된 카메라를 사용한다. 추적할 개체가 2개의 카메라에 모두 잡히도록, 카메라들은 서로 충분한 공통시야를 확보해야 한다. 그림 2의 주황색 과정들(Object Detection, 2D Tracking)은 각 카메라별로 독립 수행되며, 추적 개체들의 2차원 픽셀(pixel) 좌표를 결과로 제공한다. 반면에, 하늘색 과정들(Camera Calibration, Entity Matching, Triangulation)은 카메라들 간의 상관관계를 구하고, 그를 바탕으로 주황색 과정들의 결과인 2차원 이미지 픽셀 좌표들로부터 3차원 좌표를 계산한다.

2.2 2D Localization

개체들의 2차원 좌표 측정은 크게 object detection과 2D tracking의 두 작업으로 이루어진다. Object detection은 특정 종류의 개체들의 위치를 탐지한다. 최근 쉽게 접할 수 있는 예가 상용 디지털 카메라의 얼굴 인식 기능이다. 카메라 화면상에 얼굴로 인식되는 부분은 자동으로 사각형으로 표시가 되는 것을 한번 쬐은 본 적이 있을 것이다. Object detection에는 일반적으로 machine learning 알고리즘이 사용된다. 다양한 모습의 얼굴들을 탐지하기 위해서 얼굴의 공통적 패턴 인식이 필요하며, 여기에 machine learning 알고리즘이 사용된다. 수백 장의 다양한 얼굴 사진과 그 외의 사진들을 구분/입력하여 훈련과정을 통해 template을 생성하고, 이를 이용하여 사진 또는 비디오 프레임 내의 얼굴 부분을 인식하고 표시하는 것이다. 그림 3은 안전조끼를 입은 건설인부를 탐지하기 위한 detection 알고리즘을 나타내고 있다. 물체의 움직임, 형상, 색깔의 특성을 이용한 3단계로 이루어져 있으며, 2번째 단계와 3번째 단계의 결과(그림 3 하단, 왼쪽으로부터 3번째와 4번째)를 비교함으로써 안전조끼의 착용여부를 가려낼 수도 있다 (Park and Brilakis 2012b). 건설 장비의 경우, 보는 각도에 따라 서로 다른 외형을 보이기 때문에 전면/측면/후면별로 별도의 template을 생성할 필요가 있다 (그림 4).

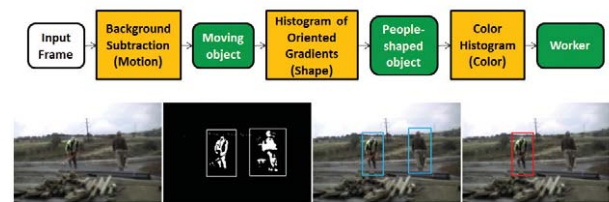


그림 3. Construction worker detection (Park and Brilakis 2012b)



그림 4. Wheel loader detection templates (Park and Brilakis 2012a)

이러한 object detection을 비디오에 적용할 경우, 프레임마다 건설 개체들의 위치를 제공하지만, 같은 종류 내에서 개체들 간의 구분은 이루어지지 않는다. m번째 프레임에 건설인부가 p_{m1} , p_{m2} 에 위치하고 있고, n번째 프레임에 건설인부가 p_{n1} , p_{n2} 에 위치하고 있다는 정보는 알 수 있지만, p_{m1} 에

위치했던 인부가 p_{n1} , p_{n2} 중 어느 위치에 해당하는지는 알 수 없다 (그림 5). 즉, 프레임마다 독립적인 위치데이터를 제공할 뿐, 프레임 간의 상관관계는 없으며, 개체별 추적은 이루어지지 않는다. 이러한 부분을 2D tracking 알고리즘을 통해 해결한다. 2D tracking은 object detection에 의해 주어진 위치를 입력받고, 각각의 개체별로 다음 프레임에서의 위치를 예측한다. Object detection과 2D tracking 알고리즘들은 동시 병행으로 실행되며, 개체가 화면에서 사라졌을 때, 자동으로 프로세스를 중단하고, 새로 출현하였을 때, 다시 그 개체를 탐지하여 추적하게 된다.

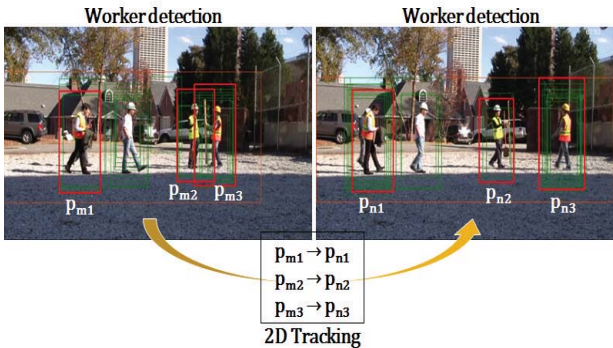


그림 5. Object detection & 2D tracking

2.3 카메라 상관관계 추정 및 3차원 좌표 계산

비디오 프레임 상의 2차원 위치 정보는 건설관리 작업에 적용하기에는 한계가 있다. 추적된 개체들이 실제로 카메라로부터 얼마나 멀리 위치하는지 깊이에 대한 정보가 없다. 보다 넓은 활용도를 위해서는 3차원 위치 정보를 확보해야 한다. 이를 위해 2대의 카메라를 이용하고, 2대의 카메라로부터 얻어진 2차원 위치 좌표들에 triangulation 알고리즘

을 적용하여 최종적으로 3차원 위치좌표를 얻는다. Triangulation을 위해서는 camera calibration이 우선되어야 한다. 이는 위치와 각도가 고정된 두 카메라들의 내부 변수들(초점거리, 왜곡변수 등)과 외부변수들(카메라들 간의 거리 및 각도)들을 찾아내는 과정이다 (그림 6). 이 작업은 카메라 설치 후, 시스템 시작 초기에 한 번만 실행된다. Triangulation의 입력 변수는 camera calibration 결과들과 2차원 위치 정보들이다. 개체별로 비디오 프레임마다 각 카메라로부터 하나씩 총 두 개의 2차원 추적 결과(p_l, p_r)가 2.2장의 과정을 통해 얻어진다. 이 두 좌표들은 최종적으로 triangulation을 통해 하나의 3차원 좌표(P)로 변환된다.

3. 실험 및 결과

제안된 시스템의 정확도를 테스트하기 위하여 실제 건설 현장에서 두 대의 카메라를 이용한 위치 추적을 시행하였다. 개체가 움직일 경로를 미리 현장에 표시하였고, total station을 이용하여 경로의 실제 위치를 먼저 측정하였다 (그림 7).

추적 개체의 이동 경로는 카메라로부터 대략 40~50m 거리에 위치하였다. 건설인력 1명과 SUV 차량 1대에 대하여 위치 추적 실험을 진행하였다. 건설인력은 지그재그 형태의 이동경로를, SUV 차량은 직선의 경로를 따라 이동하였다. 카메라 간의 거리 3.8m, 8.3m 두 경우에 대하여 실험을 하였다. 카메라 간의 거리가 8.3m인 경우에 보다 정확한 결과를 얻을 수 있었으며, 두 개체 모두 95% 신뢰도 기반, 최대 0.65m의 오차 이내로 추적되었다 (Park et al, 2012). 그림 8은 추적 결과를 보여주고 있다. 검은 실선은 total station

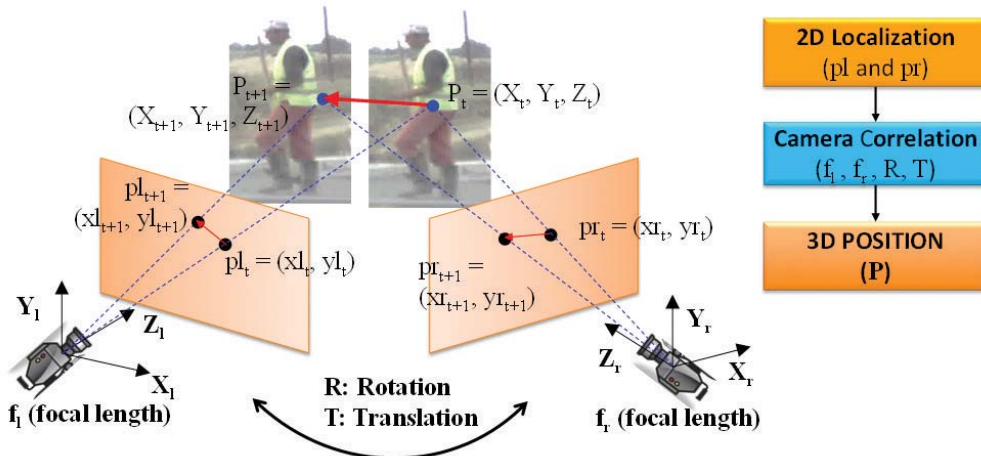


그림 6. Triangulation에 의한 3차원 위치 결정

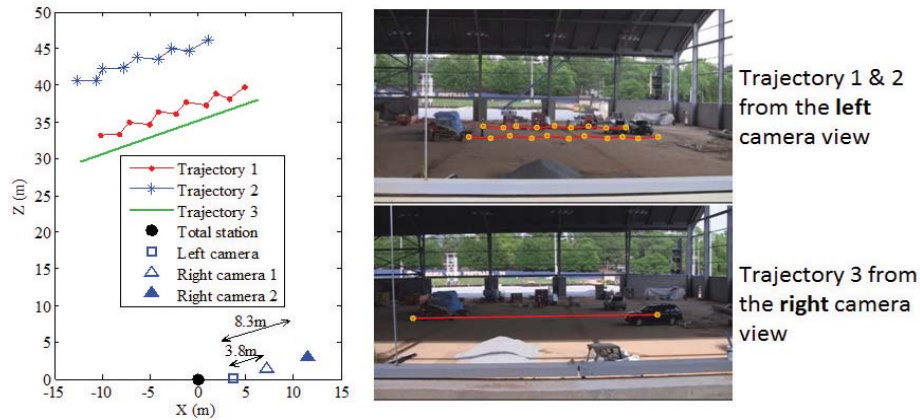


그림 7. 현장 테스트 layout (Park et al. 2012)

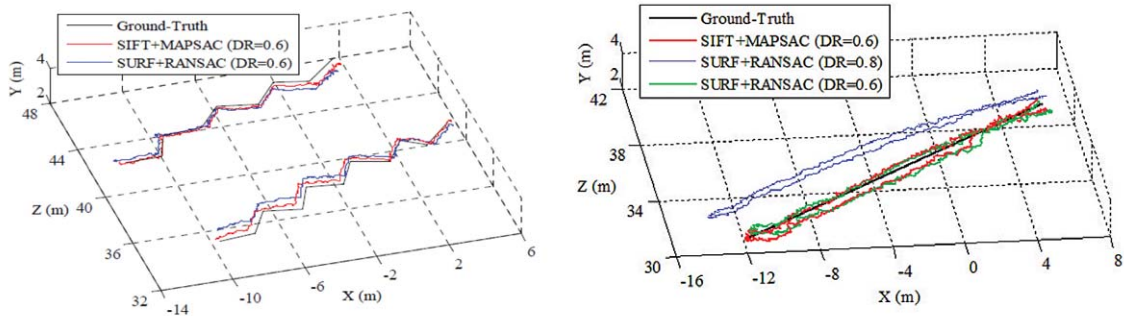


그림 8. vision-based tracking 결과 (Park et al. 2012)

의 결과(ground-truth)를 나타내고 있으며, 빨간색, 파란색, 초록색 결과는 서로 다른 camera calibration 방법을 이용한 vision-based tracking 결과이다.

4. 결론

건설현장은 인력, 장비, 자재 등 다양한 종류의 개체들이 계획된 공정에 맞추어서 시시각각 서로 다른 움직임을 보이는 공간이다. tracking 시스템은 이러한 개체들의 위치 변화를 감지하여, 현장관리에 유용한 정보를 제공해 준다. 기존의 RFID, GPS에 비해 시간 및 비용 면에서 보다 효율적인 대안이 될 수 있는 vision-based tracking 시스템을 소개하였다. 두 대의 카메라를 현장에 고정 설치하여, 카메라 화면에 잡힌 모든 개체들의 위치를 추적하는 데에 궁극적인 목적이 있다. 현장에서의 예비실험을 통해 그 정확도와 활용성을 엿볼 수 있었다. 현재, 시스템의 기반을 다지고 있는 과정에 있으며, 안정적이고 신뢰도 높은 시스템 구축을 위해서, 다양한 종류의 개체들, 다양한 환경의 현장에 대한 실험과 그를 통한 꾸준한 시스템 개선이 이어져야 할 것이다. 더 나아

가, 현장조건, 추적개체 종류 및 개체 수 등 다양한 조건에 따른 RFID, GPS, vision-based 시스템들의 적합한 조합을 제시하는 것도 중요하다고 사료된다.

참고 문헌

- Park, M.-W. and Brilakis, I. (2012a), "Enhancement of Construction Equipment Detection in Video Frames by Combining with Tracking." *Computing in Civil Engineering* (2012): pp. 421-428.
- Park, M.-W. and Brilakis, I. (2012b), "Construction worker detection in video frames for initializing vision trackers." *Automation in Construction*, 28, pp. 15-25.
- Park, M.-W., Koch, C. and Brilakis, I. (2012), "Three-dimensional tracking of construction resources using an on-site camera system." *Journal of Computing in Civil Engineering*, 26 (4), pp. 541-549.