

발 간 등 록 번 호

11-1613000-001555-01

최
중
보
고
서

건설
현장
모니터링을
위한
건설인력용
영상
기반
위치
추적
시스템
응용기술
개발

건설 현장 모니터링을 위한 건설인력용 영상 기반 위치 추적 시스템 응용기술 개발 최종보고서

2016.07.29

Infrastructure
R&D Report

주관연구기관 / 명지대학교

국
토
교
통
부
국
토
교
통
과
학
기
술
진
흥
원

주 의 사 항

1. 이 보고서는 국토교통부가 출연하고 국토교통과학기술진흥원이 위탁 시행한 국토교통기술 연구개발사업의 최종보고서입니다.
2. 본 보고서의 내용을 발표할 때에는 반드시 국토교통부가 출연하고 국토교통과학기술진흥원이 위탁 시행한 국토교통기술연구개발사업의 연구결과임을 밝혀야 합니다.
3. 국가과학기술 기밀유지에 필요한 내용은 대외적으로 발표 또는 공개하여서는 아니 됩니다.



국 토 교 통 부
국토교통과학기술진흥원

제 출 문

강호인 장관 귀하

이 보고서를 "건설 현장 모니터링을 위한 건설인력용 영상 기반 위치 추적 시스템 응용기술 개발에 관한 연구" 과제의 보고서로 제출합니다.

2016. 10. 11

주관연구기관명 : 명지대학교 산학협력단

주관연구책임자 : 박만우

연구 원 : 이용주

보고서 요약서

과제고유번호	14CTAP-C 078828-01	해 당 단 계 연 구 기 간	2014.8.14. ~2016.8.26	단 계 구 분	1/1
연 구 사 업 명	국토교통기술촉진연구사업				
연 구 과 제 명	최 상 위 과 제 명	건설 현장 모니터링을 위한 건설인력용 영상 기반 위치 추적 시스템 응용기술 개발			
	단위과제명	건설 현장 모니터링을 위한 건설인력용 영상 기반 위치 추적 시스템 응용기술 개발			
연 구 책 임 자	박만우	총연구기간 참 여 연구원수	총 : 명 내부 : 명 외부 : 명	총연구비	정부 : 천원 기업 : 천원 계 : 천원
연구기관명 및 소 속 부 서 명	명지대학교 토목환경공 학과	참여기업명			
국제공동연구	상대국명 :				

는 것을 목적으로 함

- 단일 카메라의 영상 내에서 건설인력을 탐지하고 추적하여 2차원 픽셀좌표의 시간이력 데이터를 안정적으로 획득하는 기법을 확립함
- 추가적으로 안전모를 탐지하고 건설인력과 상대적 위치를 고려하여, 건설인력의 안전모 착용여부를 판단하는 기법을 확립함
- 스테레오 영상 기반 위치추적을 위한 카메라 칼리브레이션 기법들을 비교 분석하여 효율적, 효과적 방식을 선정함
- 2대의 카메라 영상에서 탐지된 건설인력들 간의 매칭 기법을 확립함
- 3차원 위치추적이 가능한 플랫폼을 마련함
- 3차원 위치추적 결과의 정확도를 검증하기 위하여 GPS 또는 토탈스테이션 데이터와의 비교실험을 수행할 예정임

색 인 어 (각 5개 이상)	한 글	프로젝트 모니터링, 위치 추적, 건설인력, 영상 처리, 카메라
	영 어	Project Monitoring, Tracking, Construction Worker, Image Processing, Camera

요 약 문

I. 제 목

건설 현장 모니터링을 위한 건설인력용 영상 기반 위치 추적 시스템 응용기술 개발

II. 연구개발과제 개요(목적, 필요성, 범위 등을 기술)

본 연구는 건설현장에서 건설인력의 위치 추적을 위한 영상기반 시스템을 제안한다. 현장 개체들의 실시간 위치 데이터는 진척도 관리, 일정 관리, 안전 관리 등 현장 시공관리에 중요한 정보를 제공하며, 궁극적으로 효율적 현장 관리를 가능하게 한다. 무선 주파수 기반의 RFID, GPS 기술들이 현장에서의 위치 추적 용도로 사용된 바가 있으나, 추적하고자 하는 개체마다 센서(tag)를 부착해야 하는 단점이 존재한다. 반면에, 영상기반 시스템은 현장에 설치된 소수의 카메라만으로 다수의 현장 개체들을 추적할 수 있다는 장점이 있다. 즉, 시간/비용 측면에서 보다 효율적인 시스템이다. 이미지 패턴 분석과 컴퓨터 비전 기술을 적용하여 현장에서 건설인력의 3차원 좌표를 계산하는 시스템(카메라+프로세서)을 개발하고자 하며, 간단한 현장 테스트를 통하여 결과의 활용 가능성과 범위를 제시하고자 한다.

III. 연구개발 수행내용 및 결과(연구수행과정, 연구수행결과 등을 기술)

- 건설인력과 안전모 탐지 기법 개발
 - 움직임, 형상(standing, crouching), 색상 데이터를 이용한 건설인력 탐지 알고리즘 구현
 - 영상에 나타난 안전조끼 착용 인력을 99% 이상의 precision으로 탐지하여 2D tracking과의 연계를 위한 요건 충분히 확보
- 안전모 착용여부 탐지 기법 개발
 - 형상 데이터를 이용한 안전모 탐지 알고리즘 구현
 - 탐지된 안전모와 건설인력 간의 상대 위치 및 크기 비율 데이터를 기반으로 안전모와 건설인력 간의 매칭 기법 확립
 - 영상에 나타난 인력 중 안전모를 착용하지 않은 인력을 94% precision, 89% recall의 정확도로 탐지
 - 안전모 또는 안전조끼를 착용하지 않은 인력을 86.8% precision, 97.9% recall의 정확도로 탐지
- Object detection과 2D tracking의 연계 알고리즘 개발

- 건설인력 탐지 알고리즘을 이용한 2D tracking 초기화
- 개체가 화면에서 사라질 경우, 2D tracking을 자동 종료하고, 다시 새롭게 나타날 경우 재개하는 과정에 대한 파라미터 연구
- Object detection에 의한 2D tracking의 불안정성 보완 기법 확립 및 구현
- 단일 카메라를 사용하여 안전조끼를 착용한 건설인력의 연속적인 2차원 위치 추적 구현 (precision: 98.9%, recall: 98.0%)
- 다양한 occlusion(추적 중인 인력 개체가 다른 개체에 의해 가려지는 현상) 조건에 대한 안정성 확보
- 한 번 사라졌다가 다시 나타난 개체를 동일 개체로 인식하지는 못함
- Entity Matching 기법 확립
 - 동일한 장면을 서로 다른 각도에서 촬영한 2개의 영상에서 각각 탐지된 건설인력 리스트들을 쌍을 맺어주는 entity matching 알고리즘을 epipolar geometry 정보를 기반으로 구현
 - Matching 성능의 precision 향상을 위해, distance ratio thresholding, frame grouping, motion direction 비교 등 3가지 기법을 도입
 - distance ratio thresholding과 motion direction 비교 기법을 조합하여 100% precision의 매칭 결과를 얻음
- 스테레오 비전 시스템 기반의 3차원 위치 추적
 - 고정된 캠코더 카메라 2대로 이루어진 스테레오 비전 기반으로 3차원 위치 좌표 획득을 위한 카메라 칼리브레이션 방법 구현
 - 칼리브레이션 방법에 따른 3차원 위치 좌표의 정확도를 분석하기 위하여 GPS 및 토탈스테이션 데이터와의 비교 실험을 계획 (실험 진행 중)

IV. 최종 연구성과 및 적용실적

- 국제 저널 논문
 - Yang, J., **Park, M.-W.**, Vela, P.V., and Golparvar-Fard, M. (2015) "Construction performance monitoring via still images, time-lapse photos, and video streams: Now, tomorrow, and the future", Advanced Engineering Informatics, 29, 211-224.
 - **Park, M.-W.**, Elsafty, N., and Zhu, Z. (2015) "Hardhat-Wearing Detection for Enhancing On-Site Safety of Construction Workers", Journal of Construction Engineering and Management, 141(9): 04015024.
 - **Park, M.-W.** and Brilakis, I. (2016) "Continuous Localization of

Construction Workers via Integration of Detection and Tracking”,
Automation in Construction (in 2nd-round review)

- 국제 학술대회 발표

- Zhu, Z., **Park, M.-W.**, and Elsafty, N. (2015) “Automated Monitoring of Hardhats Wearing for Onsite Safety Enhancement”, Proceedings of the 5th International Construction Specialty Conference (ICSC 2015), June 7-10, Vancouver, Canada.
- **Park, M.-W.**, and Brilakis, I. (2015) “Improved Localization of Construction Workers in Video Frames by Integrating Detection Tracking”, Proceedings of the 5th International Construction Specialty Conference (ICSC 2015), June 7-10, Vancouver, Canada.
- **Park, M.-W.**, and Zhu, Z. (2015) “Vision-Based Identification of Personal Protective Equipment Wearing”, Proceedings of the 6th International Conference on Construction Engineering and Project Management (ICCEPM 2015), October 11-14, Busan, Korea.
- **Lee, Y.**, **Park, M.-W.**, and Brilakis, I. (2016) “Entity Matching across Stereo Cameras for Tracking Construction Workers”, Proceedings of the 33rd International Symposium on Automation and Robotics in Construction (ISARC 2016), July 18-21, Auburn, AL.

- Software

스테레오 영상 데이터로부터 건설인력의 위치를 추적하는 알고리즘들을 .NET Framework을 기반으로 한 C# 프로그래밍과 Matlab 프로그래밍으로 구현하였으나, 아직 하나의 플랫폼으로 통합되지는 않았으며, 소프트웨어 등록은 이뤄지지 않았음.

V. 연구목표 달성 및 효과

- 스테레오 비전 기반의 건설인력 3차원 위치추적 시스템을 구현하기 위해 필요한 핵심 요소들 중 기존에 뚜렷한 연구결과가 발표된 바 없는 다음의 요소들에 대한 새로운 알고리즘을 개발하였음
 - Object Detection과 2D tracking의 연계
 - Entity Matching
- 상기의 주요 목표 외에 탐지 가능한 건설인력의 자세를 추가하였으며(꾸그린 자세), 안전모 탐지를 추가 도입하여 안전모 착용 여부를 판단할 수 있는 영상 기반 자동화 시스템을 구현하였음
- 최종 결과에 해당하는 3차원 위치좌표 획득 알고리즘을 구현하여 정확도

등의 성능 평가를 위한 실험 중에 있음.

- 건설 현장에서 다수의 건설인력 위치데이터를 일반 카메라만으로 획득할 수 있는 영상기반 위치추적 시스템의 기반을 마련함. 현장에 바로 적용하기에는 아직 이른 상태이며, 지속적인 연구와 다양한 테스트를 통해 안정성을 확보할 필요가 있음

VI. 연구성과 활용계획

- 영상처리 및 컴퓨터비전 기술은 건설현장에서도 활용될 수 있는 잠재력이 풍부한 기술로서 오래기간 활발한 연구들이 진행되어 왔음. 하지만 위치추적 시스템으로서의 활용에 대한 연구는 아직도 미비한 상태이며, 특히 건설인력의 위치추적에 대한 연구는 찾아보기 힘든 상황임
- 현장에서의 장비 및 인력의 위치데이터는 현장상황 파악에 유용한 정보가 될 수 있으며, 실제로 위치추적 데이터의 패턴분석을 통한 현장상황 파악에 대한 연구들이 진행되고 있음
- 이 연구를 통해 확립한 안전모, 안전조끼 착용여부 자동 탐지 기술은 현장 안전관리에 효용성이 높을 것으로 사료되며, 정확도 향상을 위한 추가적 연구를 통해 현장 적용이 가능해질 것으로 사료됨
- 이 연구는 건설현장에서의 카메라 활용도를 높일 수 있고, 저비용 고효율의 건설인력 위치데이터 기술을 확보할 수 있는 기반을 다지고 있음. 지속적인 연구를 통해 안정성이 확보되면 현장에서 활용가능한 수준의 시스템 개발이 가능할 것으로 사료됨

SUMMARY

Tracking construction workers onsite yields trajectory data that can be very useful in safety and productivity monitoring. Tracking via cameras and computer vision-based methods is one of the tracking options available and deemed suitable for worker tracking in congested outdoor environments. It provides the location of a large number of entities in a camera view across frames. However, this option has yet to yield a practical, robust solution, because of its poor performance in handling occlusions and maintaining trajectories over time. Also, vision trackers provide only two-dimensional (2D) pixel coordinates, which are not adequate for construction applications. This research aims to develop and implement a novel framework of tracking multiple on-site workers in 3D using the stereo vision system that consists of two fixed cameras. The framework consists of several core functions including object detection, 2D tracking, camera calibration, entity matching, and triangulation. The workers are detected and tracked in two camera views independently through the combination of object detection and 2D tracking. In other words, each worker is tracked in 2D camera views and 2D pixel coordinates are obtained from the two views. The 2D pixel coordinates are triangulated based on the camera calibration results to calculate 3D coordinates. This research project focuses mainly on the combination of object detection and 2D tracking, and entity matching, which have not been investigated before. The final objective is to integrate all the functions into the single platform which takes videos from two cameras and provides trajectories of construction workers in 3D space. As planned, the object detection and 2D tracking algorithms are combined together to supplement each other and to track workers continuously and stably. The combination also allows to handle various occlusion cases. In the process of detecting workers, two poses are included - standing and crouching. Furthermore, the hardhat detection is implemented additionally, which enables to identify workers without wearing hardhats. This application is expected to help on-site safety monitoring. Prior to the triangulation, the entities tracked in two different camera views need to be matched across the views. Accordingly, this research proposes an entity matching method that performs with high precision. Several ways of camera calibration methods are revisited, and two methods that can be applied to the stereo vision system with long baselines are selected and implemented. Experiments are planned to evaluate the performance of the 3D tracking that integrates all required functions.

CONTENTS

Chapter 1 Summary of Research

- Research Objectives and Needs Statement	10
- Research Problem, Scope, and Specifications	13
- Current State of Research	15
- Intellectual Merit and Impact	19

Chapter 2 Research Methodology and Validation

- Research Strategies	20
- Research Methodology and Validation	21

Chapter 3 Research Outcome

- Journal Papers	61
- Conference Proceedings	61
- Software	62

Chapter 4 Research Achievements and Impact

- Achievement Level	63
- Technical and Economic Impact Analysis	64

Chapter 5 Utilization of Research Outcomes and Needs for Follow-Up Research

- Plans for Utilizing Research Outcomes	66
- Needs for Follow-Up Research	66

Reference	67
------------------------	-----------

Research Modification	70
------------------------------------	-----------

목 차

제1장 연구개발과제 개요

제1절 연구개발의 목적 및 필요성	10
제2절 연구개발의 정의, 범위, 특징	13
제3절 관련분야의 국내외 기술개발 현황	15
제4절 연구개발결과가 관련분야에서 차지하는 위치	19

제2장 연구개발 수행 내용 및 결과

제1절 연구수행 전략 및 방법론	20
제2절 연구수행과정, 내용 및 결과	21

제3장 최종 연구성과 및 적용실적

제1절 국제저널 학술지 게재 논문	61
제2절 국제 학술회의 발표	61
제3절 Software	62

제4장 연구목표 달성 및 효과

제1절 연구개발 최종목표 달성도	63
제2절 연구개발 성과의 기술적 효과분석	64
제3절 연구개발 성과의 경제적 효과분석(시장창출, 고용창출 등)	65

제5장 연구성과의 활용 및 추가연구 필요성

제1절 연구성과 향후 활용방안(기술실시계획, 판매계획, 사업화 추진방안, 신기술 신청, 정책반영 계획 등, 기존 연구 및 타연구와의 연계 등)	66
제2절 추가연구 필요성 등	66

참고문헌	67
------------	----

중요 연구 변경사항	70
------------------	----

제1장 연구개발과제 개요

제1절 연구개발의 목적 및 필요성

1. 연구개발 배경 및 목적

3차원 위치 추적 시스템은 움직이는 개체들을 추적하여 개체들의 3차원 위치 데이터를 제공할 수 있다. 간단한 데이터 후처리 과정을 통해, 실시간 위치 데이터로부터 각 개체들의 이동 방향, 이동 시간, 이동 패턴을 계산하는 것 또한 가능하다. 개체들의 위치 및 이동과 관련된 정보들은 건설현장의 시공 상황을 엿볼 수 있는 중요한 모니터링 자료로 활용이 가능하다. 선별된 현장 모니터링 자료는 궁극적으로 공사 과정 현황을 분석하고 필요에 따라 신속하고 정확한 결정을 내릴 수 있는 근거가 된다. 이러한 다양한 활용성 때문에, 건설현장에서 특정 개체들의 3차원 위치 추적을 위한 연구들이 꾸준히 진행되어 왔고, 주로 무선 주파수 기술을 기반으로 한 RFID(Radio Frequency Identification), GPS(Global Positioning System), UWB(Ultra Wide-Band) 기반 방식들이 주목을 받아왔다. 많은 연구 결과들을 통해 이러한 시스템들의 실제 활용가능성을 엿볼 수 있었으나, 무선 주파수 기반 방식들은 추적하고자 하는 개체마다 센서나 태그 등을 부착해야만 한다. 이러한 특징으로 인해 수많은 장비와 인력으로 밀집된 건설현장에 적용하기에는 한계가 있으며 추적하고자 하는 개체가 많아질수록 비용이 증가하게 된다.

이 연구에서는 이러한 점에 착안하여, 이에 대한 대안으로서 센서나 태그의 부착 없이도 개체를 추적할 수 있는 영상기반 위치 추적 시스템을 제안하며, 추적 대상을 건설인력으로 좁혀 그 활용 가능성을 타진하고자 한다. [근로자 참여 및 협력증진에 관한 법률] 제 20조에 의하면, 사업장 내 근로자 감시 설비의 설치에 노동자들과의 협의를 필요로 하는 사항이다. 따라서, 카메라 설치 및 정보수집의 목적이 산업 안전 및 시공관리에 있음을 분명히 하여, 원만한 협의를 이루어야 한다. 영상 속의 인력 추적을 통해서도 신분을 인식하기가 어렵기 때문에, 인권 침해의 여지가 거의 없으며, 기록되는 위치 데이터에 신분에 대한 정보가 배제됨을 확인한다면 인권 침해에 대한 논쟁은 없을 것으로 판단된다.

2. 건설 현장에서의 위치추적 데이터 활용성

위치추적 시스템은 추적 개체들의 위치를 시간이력으로 제공한다. 일반적인 건설 현장의

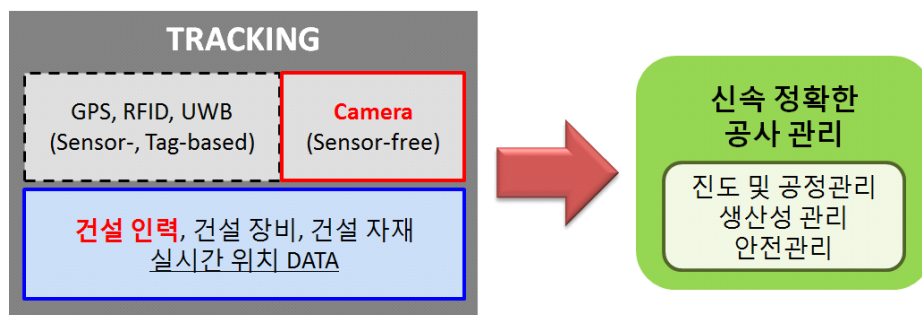


그림 1-1. 연구 목적 개요

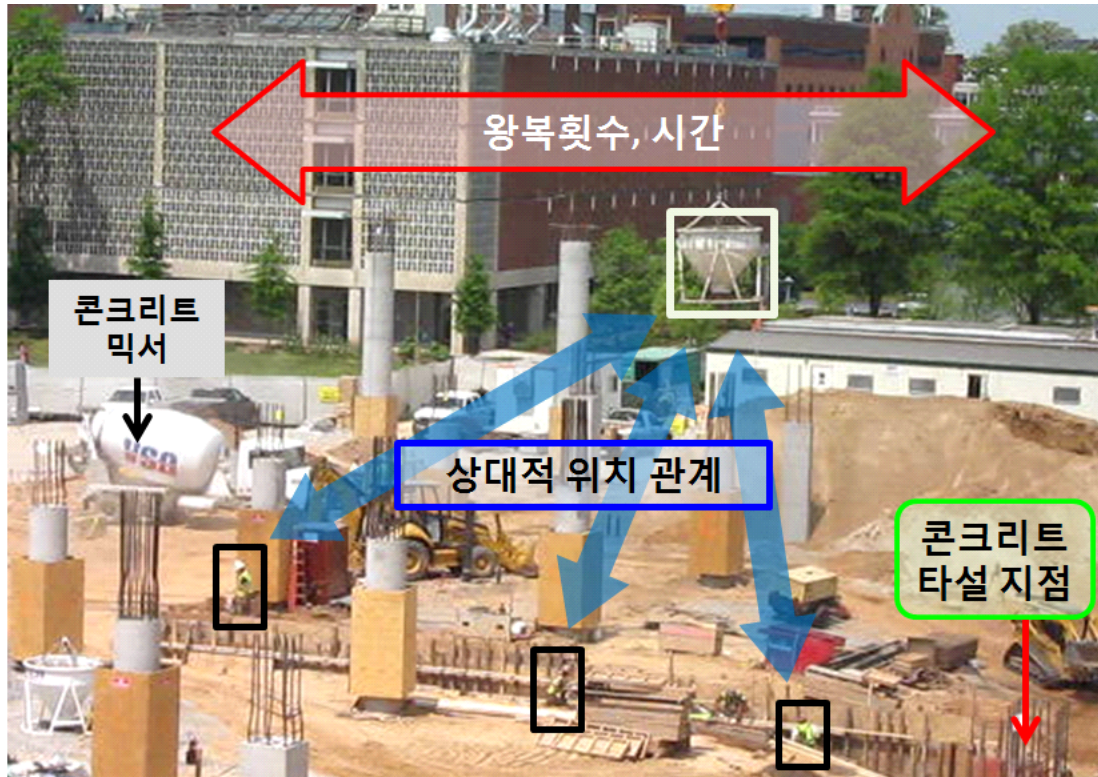


그림 1-2. Tracking의 활용도 예시

복잡성, 밀집도, 그리고 다양한 개체들의 움직임을 고려해 볼 때, tracking 시스템이 제공하는 데이터는 효율적 건설 관리 및 현장 모니터링에 중요한 정보로 활용될 수 있다. 건설인력들과 다양한 종류의 건설장비들, 그리고 건설 자재들의 위치를 실시간으로 취할 수 있다면, 이는 진도 관리, 현장 안전 관리, 생산성 분석, 공정 관리 과정의 신뢰도를 높이고 체계적 관리를 이루는 데에 이바지할 수 있을 것이다. 그림 1-2는 다수의 콘크리트 기둥이 시공되고 있는 빌딩건설 현장이다. 콘크리트 기둥의 현장 타설을 위해, 콘크리트 버킷은 콘크리트 믹서와 콘크리트가 타설되어야 할 기둥지점 사이를 타워크레인에 매달려 수차례 왕복하게 된다. 이러한 현장에서, 콘크리트 버킷을 추적한다면, 기둥 1개를 타설하기 위해 소요되는 콘크리트 버킷의 왕복 횟수와 왕복 시간을 구할 수 있으며, 이는 생산성 관리에 중요한 데이터로 활용될 수 있다. 또한, 어느 지점의 기둥이 타설되고 있는지를 바로 파악하여, 진도 및 공정관리의 효율성을 향상시킬 수 있다. 여기에 추가적으로 건설인력의 추적데이터가 더해진다면, 이동하는 콘크리트 버킷과 건설인력의 상대적 위치 관계를 파악할 수 있으며, 이를 바탕으로 콘크리트 버킷 밑으로 인력이 움직이지 않도록 미리 경고 메시지를 주어서 안전관리를 위한 시스템으로도 활용 가능하다.

3. 영상 기반 Tracking 시스템의 활용 가능성

가. 그림 1-3에 요약되어 있듯이, GPS, RFID, UWB 등은 주로 건설 장비와 자재의 추적으로 많이 연구되어 왔으며, 센서나 태그를 개체마다 부착해야하기 때문에 추적 개



그림 1-3. 영상 기반 Tracking 시스템의 활용 가능성

체가 늘어남에 따라 그 효율성과 적용성이 저하된다. 이로 인해, 센서나 태그를 필요로 하지 않는 tracking 시스템이 요구된다. 카메라를 사용하는 영상 기반 시스템은 sensor-free 시스템으로서 무선 주파수 기반 시스템의 단점을 보완할 수 있으며, 특히 센서 부착이 자유롭지 않은 건설인력의 추적에 적합할 것으로 기대된다.

나. 영상 기반 추적은 센서나 태그를 필요로 하지 않고, 스포츠 중계에서 선수들이나 공의 위치를 측정하는 데에 널리 이용되고 있다.

다. 넓고 밀집된 현장에서의 위치 추적을 위한, 보다 효율적인 대안으로서의 잠재성을 갖고 있다. 그림 1-4에서 나타냈듯이, 축구장을 건설 현장으로, 축구 선수들을 건설 장비나 인력으로 대체해 봄으로써, 영상 기반 추적 시스템의 적용 가능성을 간접적으로 엿볼 수 있다.

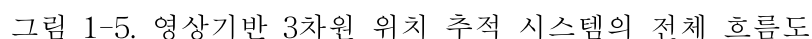
라. 컴퓨터 비전 분야의 기술들이 진화되고, 카메라 및 CPU의 성능들이 지속적으로 발전하면서 영상 기반 추적 시스템의 적용성과 가능성은 더욱 부각되고 있다. 이러한 영상 기반 추적 기술을 건설 현장에서의 위치 추적 시스템에 적용한다면, 무선 주파수 기반 시스템들의 단점들을 보완할 수 있을 것이다.



그림 1-4 Vision-based tracking의 적용성

- ## 제2절 연구개발의 정의, 범위, 특징

라. 실제 현장의 영상을 이용한 예비 테스트를 실시하여, 정확도, 활용 가능성, 필요 개선 사항을 분석하여 추후 연구 방향을 정립한다.



마. GPS, RFID, UWB 등의 무선주파수 기술에 대한 대안으로서 영상기반 시스템이 갖는 장점 중의 하나는 저비용의 장비 사용이다. 카메라 2개와 영상 프로세싱을 컴퓨터 장비만을 필요로 하므로, 비용 면에서 효율적인 tracking 시스템을 구축하는 것을 이 연구의 궁극적인 목표이다.

2. 연구개발의 특징

가. 본 연구과제의 궁극적 목표는 그림 1-5에 표현된 건설인력 추적을 위한 영상기반 시스템의 framework을 하나의 소프트웨어로 구현하는 것이다. 소프트웨어의 각 구성요소가 될 그림 1-5의 각 단계들의 관계를 정의하고, 기존의 연구가 미비한 단계들에 대해 심층적으로 분석하여 알고리즘들을 확립한다. 따라서, 본 연구는 기술의 개념 및 문헌 연구 수준을 넘어 관련 개념을 기술적으로 구현하기 위한 실험적 연구에 해당한다.

나. 그림 1-5의 주요 요소들인 camera calibration, object detection, 2D tracking, 삼각측량(triangulation) 등은 모두 이미지 프로세싱, 컴퓨터 비전 분야에서 활발한 연구가 진행되어 온 분야이며, 이 연구에서는 새로운 알고리즘을 개발하기보다는, 실험 및 검증을 통하여 최적의 알고리즘을 선별하고, 올바른 적용과정을 정립하는 데에 초점을 둔다.

다. 컴퓨터 비전 기술은 빠른 속도로 발전하고 있으며, 다양한 알고리즘들이 개발되고 있다. 이 중에는 건설분야에서도 활용 가능한 잠재적 기술들이 존재한다. 하지만, 모든 알고리즘들이 건설분야에 바로 적용 가능한 것이 아니라, 활용도에 맞도록 일정 수준의 변경/조절 작업이 요구되며, 실제 활용 가능성에 대한 검토를 위한 별도의 검증 작업 역시 필요하다.

라. 예를 들어, 건설 인력의 탐지 및 추적을 위해서는, 건설 인력이 다른 개체와 구분될 수 있는 이미지 특성들에 대한 검토가 필요하다. 또한, stereo vision에 의한 3차원 위치 측정도 건설 현장과 같은 대규모 현장에의 적용을 위해서는, 적절한 매개변수 선정에 대한 연구가 필요하다.

마. 이 연구에서는 기존의 알고리즘들을 활용하되, 건설 인력 추적을 위한 적절한 알고리즘 선정 및 조합, 선정된 알고리즘들의 설정 및 그 검증에 대한 연구가 진행될 것이다.

바. 국제 협력: 해외에서는 이미 많은 건설공학 연구자들이 stereo-vision을 비롯한 컴퓨터 비전 기술 및 이미지 프로세싱 기법의 활용방안에 대한 연구들을 진행하고 있다. 책임연구자는 다음의 해외 연구자들과의 협력을 지속적으로 유지하면서, 현장 데이터 공유, 알고리즘 검토, 논문 작성 과정의 질적 향상을 도모하고자 한다.

(1) Ioannis Brilakis, University of Cambridge, 영국

(2) Fei Dai, West Virginia University, 미국

(3) Zhenhua Zhu, Concordia University, 캐나다

제3절 관련분야의 국내외 기술개발 현황

1. 국가 연구개발사업 유사연구 실적

국가R&D사업관리서비스 홈페이지(www.ntis.go.kr)의 유사과제 검색 시스템을 이용하여 조회한 결과, 본 연구내용과 유사한 과제는 없는 것으로 나타났다. 많은 건설 현장에서는 이미 카메라를 설치하여 현장 모니터링에 사용하고 있으나, 방대한 양의 영상 데이터로부터 얻을 수 있는 중요 정보들이 간과되고 있는 것이 현실이다. 영상 신호 처리를 통해 보다 많은 정보들을 자동 추출하는 시스템이 구축된다면 영상 데이터의 가치를 높이고 효율적 현장관리를 도모할 수 있다. 하지만, 현재 국내에서는 이러한 연구활동이 비교적 미흡한 편이다. 이 연구는 영상 데이터 활용의 한 방편으로서 위치추적 시스템을 제안하고자 한다.

2. 무선 주파수 기반 Tracking 시스템

이러한 활용성 때문에, 건설현장에서의 tracking에 관한 연구가 꾸준히 진행되었고, 주로 무선 주파수 기반의 RFID(Radio Frequency Identification), GPS(Global Positioning System), UWB(Ultra Wide-Band) 등이 주목을 받아왔다. 최근에는 Wi-Fi나 Beacon의 활용 또한 연구되고 있다.

가. GPS를 이용한 Tracking(Oloufa외 2003, Lu외 2007)

- (1) 인공위성, 센서, 그리고 중앙관리 모듈로 구성되며, 모바일 네트워크를 통해 센서들의 위치가 모두 중앙모듈로 전해져, 중장비들의 위치를 하나의 지도 위에 얹혀서 나타낼 수 있음.
- (2) GPS는 실외에서만 사용 가능하며, 주로 건설 중장비의 추적에 적용됨.
- (3) 저비용의 일반 GPS는 약 10m 정도의 오차를 보임(Caldas외 2004).
- (4) 높은 정확도를 위해서는 kinematic GPS의 사용이 필연적이지만, 높은 비용으로 인해 많은 수의 개체를 추적하기에는 비경제적임.
- (5) 중앙 모듈과 센서들의 설치에 소요되는 시간과 비용 또한 단점으로 지적됨.

나. RFID를 이용한 Tracking

- (1) RFID 기반의 추적 시스템은 RFID 태그, 리더, 데이터 관리 시스템으로 구성되어 있으며, 리더와 태그가 일정 거리 이내로 근접하면, 태그의 ID가 리더로 전송되어 전송된 위치와 시간이 저장됨.
- (2) 자재들이 현장에 조달되고 설치되는 과정을 모니터링하는 목적으로 사용됨.
- (3) 이러한 시스템을 품질 관리, 건설 중장비들의 충돌사고 방지 등에 적용한 연구들이 보고되었다(Wang 2008, Chae외 2010).
- (4) RFID는 연속적인 위치 추적이 아니라, 리더가 태그에 근접해 있을 때에만 그 대략적인 지역을 제공함.
- (5) 최근 RFID를 이용하여 3차원 좌표를 구하는 연구 또한 진행되었으나, 근거리에만

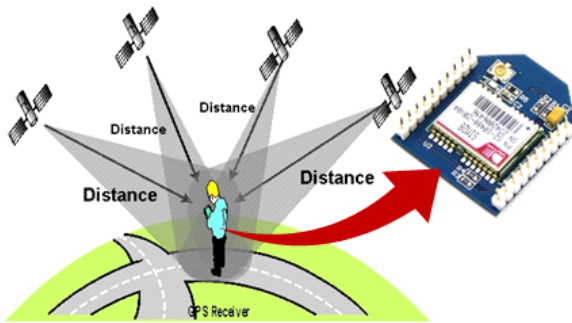


그림 1-6. 개체별 tagging이 필요한 무선주파수 기반 시스템

적용이 가능하다는 단점이 남아 있음(Ko와 2010).

- (6) RFID의 단점을 보완하기 위하여, GPS와의 통합시스템도 연구된 바 있음(Ergen와 2007).
- (7) Costin와 2인(2012)은 건설현장 안전관리 개선을 위하여 RFID를 활용한 건설 인력 추적에 관한 연구를 발표하였다. 하지만, RFID는 연속적인 위치 데이터를 제공하지 않기 때문에, 그 적용 범위가 매우 한정적이다.

다. Wi-Fi를 이용한 Tracking

- (1) 실내에서의 측위를 위해 Wi-Fi 기술을 도입하였으며, 실제 터널 시공 현장에서의 실험 결과 5m 이내의 정확도로 측위가 가능했음(Woo와, 2011).

라. 무선 주파수 기반 시스템들의 공통적 단점: 추적하고자 하는 개체마다 센서(sensor)나 태그(tag)를 부착하여야 함(그림 1-6).

- (1) 추적 개체 수가 늘어남에 따라 소요비용이 늘어나게 되며, 정해진 기간마다 센서들을 회수하고, 다시 부착하는 번거로움을 감수해야 함.
- (2) GPS가 이미 장착되어 있는 최신장비들이 출시되고 있기는 하지만, 실제 현장에서 사용되는 최신장비의 수는 매우 낮은 비율을 차지할 뿐이며, 대부분의 장비들이 GPS가 없는 장비들인 것이 현실임.
- (3) 현장 인부들을 추적하기 위해 센서를 부착할 경우 강한 반발을 야기할 수 있으며, 인부들이 센서를 임의로 제거할 경우에도 규정 상 달리 취할 방도가 없음.
- (4) 대규모의 건설장비 및 인력으로 밀집된, 넓은 규모의 현장에서는 이러한 기술들의 적용성이 현저히 떨어짐.

⇒ 추적 가능한 개체의 수를 최대화 하고, 필요 인력을 최소화하며, 시간 및 비용 면에서 효율적인 tracking 기술이 요구됨

3. 영상 기반 Tracking 시스템에 관한 연구

가. 영상 기반 2D Tracking에 관한 연구

- (1) 현장 안전관리를 목적으로 건설인력을 추적하는 연구결과가 처음 발표되었음 (Yang외 2009).
- (2) 타워크레인에 의해 운반되는 콘크리트 버킷을 추적하여 생산성 산출에 적용한 연구가 발표되었음(Gong외 2010).
 - ⇒ 단일 카메라를 사용하였기 때문에, 2차원 pixel 좌표만으로 추적하여, 공간적 이동 상황을 정확하게 파악하기 어려움.
 - ⇒ Tracking이 시작되려면 사용자가 추적할 개체를 처음에 직접 화면상에 지정해 주어야 함.

나. Object Detection과 2D Tracking의 연계 알고리즘 개발

- (1) Object Detection은 기본적으로 각각의 프레임마다 화면상에 나타난 모든 건설인력들의 위치를 제공한다. 하지만, 프레임 간에 연계가 전혀 이루어지지 않은 프레임별 독립적 결과만을 제공할 뿐이다. 따라서, 각 인력 개체들 간의 구분이 이루어지지 않으며, 따라서 개체별 위치의 시간이력 데이터는 취할 수 없다. 개체별 추적, 즉 위치의 시간이력 데이터를 제공하기 위해서는 프레임과 프레임을 연관시켜주는 2D Tracking 과정이 필요하다. 이와 관련된 연구가 진행된 바가 있으나, 일부 건설 장비에만 한정된 연구였으며, 최대 2개의 개체만을 추적하는 결과만이 보고되었다(Rezazadeh Azar외 2013). 또한, 2 개체마저도 서로 다른 종류의 장비(덤프트럭과 백호우)이어서, 둘 사이의 구분이 명확한 시나리오에만 적용 가능하다.

다. AprilTag 활용 및 Pan-Tilt-Zoom 카메라 활용

- (1) 건설장비를 탐지함에 있어서 장비의 형상이나 색상을 사용하지 않고, 특정 표식(AprilTag)을 장비에 부착하여 표식을 탐지하는 방법을 사용함(Rezazadeh Azar외, 2015a). QR 코드와 유사한 AprilTag는 탐지 정확도가 매우 높기 때문에 신뢰성이 높은 반면, 탐지를 위해서는 충분한 해상도로 표식이 화면에 잡혀야 함. 또한 부착된 표식이 카메라 화면에 잡히도록 각도가 잘 조절되어야만 함.
- (2) 위의 단점을 일부 보완하기 위해 Pan-Tilt-Zoom 카메라를 활용하였음(Rezazadeh Azar, 2015b). 장비의 형상을 탐지한 후, Zoom-in 하여 표식을 충분한 해상도로 탐지하는 방식을 사용함.

라. 영상 기반 3D Tracking에 관한 연구: 본 연구의 책임연구자가 소속되어 있던 조지아 공대의 연구실에서, 2개의 카메라를 사용하는 stereo-vision 기술을 적용하여 3차원 좌표를 계산해 주는 3차원 위치 추적 시스템을 제안하였고, 제안된 framework의 일부 단계별 연구 결과들이 차례대로 발표되었음(Brilakis외 2011a, Park외 2011, Park외 2012a, Park외 2012b).

- (1) 가장 활발한 연구들이 진행된 단계는 object detection 부분으로 건설인력과 다양한 종류의 장비들의 탐지 기법들이 연구되었다. 하지만, 다수의 개체들을 동시 자

동 추적하기 위해서는 추가적으로 도입되어야 할 알고리즘 및 단계들이 있으며, 실제 복잡한 현장의 영상에 대한 검증은 제대로 이루어진 바가 없음.

(2) Object detection과 2D tracking의 연동에 관한 부분도 보완하고 실험해야할 사항들이 남아있음. 특히, 건설인력을 대상으로 한 연구결과는 아직 많이 미비한 상황임.

(3) Entity matching에 대한 연구결과는 아직 발표된 바가 없음.

⇒ 본 연구에서는 아직 검증되지 않은 요소들에 대한 연구를 진행하고, 모든 단계별 요소들을 통합하여 단일 패키지 소프트웨어를 개발한 후, 건설현장 적용에 대한 미비한 점들을 보완하여 그 활용성 및 적용성을 높이고자 함.

4. 스테레오 비전 기술

가. 이 연구에서는 영상기반 tracking 시스템을 제안하고 있으며, 이 시스템 구성의 핵심 요소가 stereo vision이다. stereo vision은 컴퓨터 비전, 로봇틱스 등의 다양한 분야에서 활발하게 사용되고 있다. 하지만, 두 카메라 간의 거리가 짧은 경우가 대부분으로, 소규모 3차원 공간 정보 습득에 사용되고 있음.

나. 건설 분야에서의 stereo vision 활용

(1) 지형 정보를 얻기 위한 항공사진 측량에 사용됨.

(2) 구조물의 3차원 모델링 자동화에 stereo vision 기술 적용됨(Golparvar-Fard와 2009, Brilakis와 2011b).

⇒ 이동하는 카메라를 이용하며, 고정된 물체의 3차원 형상 정보를 얻기 위한 용도로 사용됨.

다. 컴퓨터 비전 알고리즘: 이 연구에서는 기존에 발표된 컴퓨터 비전, 이미지 프로세싱 알고리즘들을 활용함. 프로그래밍에는 free source인 이미지 프로세싱 라이브러리들을 사용하며, 대표적인 라이브러리가 OpenCV(<http://opencv.org/>)임.

5. 국내외 시장분석 및 산업동향

가. 국내 건설분야에서의 컴퓨터 비전 기술 활용도는 아직 미비하며, 개척되지 못한 상태이다. 미국 및 유럽의 건설업체들은 시간과 비용 모든 면에서 효율적인 영상 또는 사진 기반의 센싱 기술에 대한 관심을 크게 보이고 있으며, 소규모 전문건설 업체들로부터 대규모 일반건설 업체들에 이르기까지 다양한 업체들이 컴퓨터 비전 적용 기술에 대한 연구를 지원/진행하고 있다.

나. 기술 성숙도가 보다 높아짐에 따라, 현장 데이터 수집과 현장 경험 기반의 피드백수렴이 가능하도록 국내외 업체와의 공동연구 체제를 갖출 필요성이 생길 것이다.

다. 국외 연구기관과의 협력, 그리고 이를 바탕으로 한 해외업체와의 연계과정은 향후 궁극적으로 해외시장 진출의 가능성을 높여줄 것이다.

제4절 연구개발결과가 관련분야에서 차지하는 위치

영상처리 및 컴퓨터비전 기술은 건설현장에서도 활용될 수 있는 잠재력이 풍부한 기술로서 오래기간 활발한 연구들이 진행되어 왔다. 하지만 위치추적 시스템으로서의 활용에 대한 연구는 아직도 미비한 상태이며, 특히 건설인력의 위치추적에 대한 연구는 찾아보기 힘든 상황이다. 위치추적 시스템으로는 GPS와 RFID가 전통적으로 많은 신뢰를 받고 있는 기술이며, 최근 실내에서의 위치추적을 위해 Wi-Fi, Beacon 등의 기술들이 도입되고 있다. 카메라 영상 기반의 위치추적 기술에 대한 연구는 그에 비해 많은 연구가 이뤄지지 않는 편이다.

캐나다 Lakehead University의 Rezazadeh Azar 교수가 건설장비 탐지 및 위치추적에 관한 연구를 꾸준히 진행하고 있다. 덤프트럭 및 굴착기 등의 장비들을 영상 내에서 탐지하고 추적하는 내용은 이 연구에서 진행된 내용들과 유사한 틀 내에서 이뤄지고 있다. 다만, 건설장비들은 건설인력과 달리 카메라 위치 및 각도에 따른 건설장비의 형상이 다양하다는 점에서 개체 탐지 과정에서 보다 다양한 템플릿을 요한다는 차이점이 있다. 반면에 건설장비는 건설인력에 비해 그 숫자가 적으며, 특히 같은 종류의 장비가 한 영상 내에 여러 대 나타나는 일은 많지 않다는 점에서 위치추적이 비교적 용이하다. 건설인력은 안전조끼 및 안전모를 착용한 상태에서의 형상과 색상이 개체 간에 거의 유사하며, 건설장비에 가려지거나 인력 간에 서로 가려지는 상황이 자주 발생한다. 따라서, 건설장비의 위치추적과는 달리 보다 세밀한 occlusion handling 기술이 요구된다. 이 연구에서는 이러한 문제들을 해결하기 위한 연구들이 진행되었고, 높은 수준의 정확도를 확보할 수 있었다. 따라서, 영상 기반의 건설인력 위치추적에 대해서는 독보적인 수준의 기술 개발이 이뤄졌다고 사료된다.

기존의 영상기반 위치추적 관련 연구들이 단일 카메라에 의존해온 반면에, 이 연구는 두 개의 카메라로 이뤄진 스테레오 비전 시스템을 활용한다는 점에서 차별성을 갖고 있다. 앞서 언급된 캐나다의 Lakehead University에서의 연구는 최근 pan-tilt-zoom 기능을 갖춘 단일 카메라를 활용하여 건설장비를 탐지하고 3차원 위치좌표까지 획득하는 기술을 개발하고 있다. 하지만, 이는 건설장비에 AprilTag라는 표식을 부착하여 개체를 탐지하는 방식을 활용하고 있으며, 고정된 위치에서 작동 중인 장비들에만 적용이 가능한 것으로 사료된다. 장비에 비해 비교적 그 크기가 작고 지속적으로 위치가 변경되는 많은 수의 인력에 적용하기에는 한계가 있다. 따라서, 건설인력의 3차원 위치추적에는 이 연구에서 도입하고 있는 스테레오 비전 기술이 가장 적합하며, 이와 관련된 유사 연구는 진행된 바 없다. 따라서 이 연구는 관련분야에서 선도적인 위치에 있다고 사료된다.

제2장 연구개발 수행 내용 및 결과

제1절 연구수행 전략 및 방법론

연구범위	연구수행방법 (이론적 · 실험적 접근 방법)
건설 인력 및 안전모 탐지	• 필요 장비(카메라) 및 소프트웨어(MATLAB) 구입 • 건설 인력 및 안전모 탐지를 위한 training image 수집 • Standing, bending 두 가지 자세와 안전모에 대한 형상 탐지 템플릿 생성 • 건설 인력과 안전모 간의 매칭 기법 확립 • 성능 정의 및 평가 방식 결정 • 성능 평가 및 검증 용 영상 수집 • 실험 및 검증을 통한 최적 파라미터 선정, 보완점 개선
Detection과 2D tracking의 연계	• 건설인력 개체가 화면에서 사라짐을 탐지할 수 있는 요소에 대한 이론적 연구 • 추적되고 있는 개체들 간의 겹침 현상으로 인해 한 개체가 화면에서 가려지는 상황에 대한 처리 기법 확립 • 2D tracking의 불안정성 인식을 위한 데이터 분석 • Detection에 의한 2D tracking의 불안정성 보완 • 성능 정의 및 평가 방식 결정 • 성능평가 및 검증 용 영상 수집 • 실험 및 검증을 통한 최적 파라미터 선정, 보완점 개선
Epipolar geometry 기반 entity matching	• 문헌 조사 • epipolar geometry 기반의 매칭 기법 구현 및 한계점 파악 • Precision 최대화를 위한 추가 방안 확립, 구현, 평가 • 개체의 이미지 특성의 활용 여부 결정 • 실험 및 검증을 통한 최적 파라미터 선정
3차원 위치추적 성능 평가	• camera calibration 및 triangulation 관련 문헌 조사 • camera calibration 기법들 구현 및 비교 평가 • triangulation 기법 구현 • GPS 또는 토달스테이션 데이터와의 비교를 통한 3차원 위치추적 정확도 검증 • 최적의 camera calibration 기법 선정

제2절 연구수행과정, 내용 및 결과

1. 개체 탐지

가. 건설 인력 탐지

(1) Standing pose

건설인력 탐지는 책임연구자가 이전에 발표하였던 논문 “Construction worker detection in video frames for initializing vision trackers(Automation in Construction)”를 토대로 구현하여 사용하였다. 건설인력을 안전조끼를 입은 사람으로 정의하였고, 안전조끼를 입은 인력을 단일 카메라 영상 내에서 탐지하고, 탐지된 영역을 2차원 픽셀 좌표로 표현하여 저장한다. 움직임, 형상, 색상 등 세 가지 이미지 특성을 사용하여 건설인력을 탐지한다. 대부분이 기존의 연구결과를 답습하여 구현한 내용이므로 여기서는 간략하게만 정리하고자 한다.

이 연구에서는 고정된 카메라를 사용하였다. 따라서 움직이는 물체에 해당하는 영역의 픽셀 값만이 크게 변하고 그 외의 영역들은 비교적 일정한 픽셀 값들을 유지한다. 이러한 사실에 근거하여 움직이는 물체가 있는 곳의 영역을 구하는 방법이 그림 2-1에서의 첫 번째 단계인 Background Subtraction(Mcfarlane외 1995)이다. 움직이는 영역 내에서 사람의 형상을 나타내는 부분을 탐지하는 것이 두 번째 단계이다. 형상을 나타내는 이미지 특성으로는 HOG(Histogram of Oriented Gradients)(Dalal외 2005)를 사용하였다. SVM(Support Vector Machine) 기계학습 알고리즘을 이용하여 하나의 템플릿으로 나타내었다. 두 번째 단계가 완료되면 영상으로부터 인력 개체를 탐지한 결과를 얻게 된다.

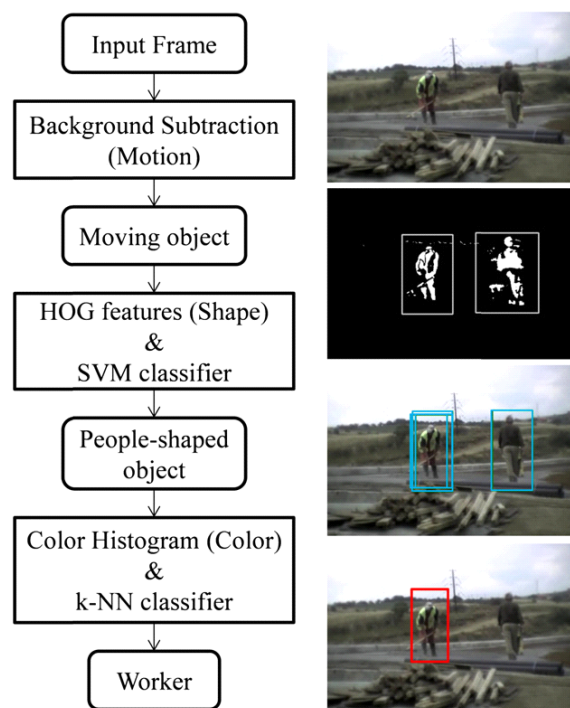


그림 2-1 건설인력 탐지 기법 구성도



그림 2-2 건설인력 탐지 결과

마지막인 세 번째 단계는 탐지된 사람의 영역에 대한 색상 정보를 추출하여 안전조끼의 착용여부를 판단/구분하는 것이다. 색상을 나타내는 특성으로는 color histogram(Swain의 1991)을 사용하였으며 RGB가 아닌 HSV(Hue, Saturation, Value)를 사용하였다. 조도, 빛의 영향을 줄이기 위하여 밝기에 해당하는 Value는 제외하였고, hue와 saturation으로 이루어진 color histogram을 구성하여 안전조끼 착용여부를 구분하도록 하였다. 다양한 안전조끼의 색상들을 모두 고려하기 위하여 수백 개의 안전조끼 사진을 훈련 데이터로 수집하였고 k-nearest neighbor 기계학습 알고리즘을 사용하였다. 결과적으로 세 번째 단계에서는 안전조끼를 착용한 사람을 탐지한 결과를 얻을 수 있으며, 두 번째 단계에서의 결과와 비교를 통해 안전조끼를 착용하지 않은 사람의 탐지 결과도 함께 얻을 수 있게 된다. 그림 2-2는 제안된 알고리즘을 사용한 결과의 한 표본이다. 초록색은 안전조끼를 착용하지 않은 사람의 영역을 나타내고, 빨간색은 안전조끼를 착용한 사람의 영역을 나타내고 있다. 개체 탐지의 성능은 일반적으로 precision과 recall로 평가한다. 건설인력 탐지에서의 precision과 recall은 다음과 같이 정의된다.

- $\text{precision} = \text{TP} / (\text{TP} + \text{FP})$
- $\text{recall} = \text{TP} / (\text{TP} + \text{FN})$
- TP(True Positive) = 탐지된 건설인력 개체 중 실제로 건설인력인 개체 수
- FP(False Positive) = 탐지된 건설인력 개체 중 실제로 건설인력이 아닌 개체 수
- FN(False Negative) = 영상에 나타난 건설인력 개체들 중 탐지되지 않은 개체 수
- 건설인력: 안전조끼를 착용한 사람

Precision과 recall은 SVM의 threshold 값에 따라서 변동하며, 일반적으로 하나가 높아지면 다른 하나가 낮아지는 상호관계를 갖는다. 이 연구에서의 건설인력 탐지 알고리즘의 주목적은 2차원 tracking 알고리즘과의 연계를 통한 정확하고 안정적인 2차원 위치추적이다. 따라서, recall 보다는 precision이 더 중요한 성능으로 고려되어 precision을 최대화하는 방향으로 SVM의 threshold를 조정하였고, 99.7% precision, 83.0% recall의 결과를 얻었다. Recall이 83.0%라는 것은 17.0%의 인력이 탐지되지 않았음을 의미하지만, 이 17.0%는 tracking 알고리즘과의 연계로 상당 부분 해결이 될 것이기 때문에 이 단계에서는 precision에 초점을 맞추어 결과를 산출하였다.



그림 2-3. Positive training images



그림 2-4. Negative training images



그림 2-5. HOG feature 예시
(좌측: standing, 우측: crouching)

(2) Crouching pose

서 있는 자세 외에 쭈그린 자세의 인력을 탐지하기 위해 추가적으로 HOG 템플릿을 생성하였다. 그림 2-3과 2-4는 HOG 템플릿 생성에 사용된 training image 샘플들이다. 그림 2-5의 좌측과 우측은 각각 선 자세와 쭈그린 자세의 HOG feature 예를 보이고 있다. 선 자세의 training image들은 64×128 해상도를 갖는 반면, 쭈그린 자세의 training image들은 64×96의 해상도를 갖도록 하였다. HOG feature 벡터의 크기는 각각 3,780와 2,772가 된다. Training image로는 718개의 positive image와 2,872개의 negative image가 사용되었다. Positive image에는 쭈그린 사람을 좌측과 우측에서 바라본 사진들이 반반씩 포함되었다.

테스트에 사용하기 위한 영상을 캠코더로 촬영하여 탐지 성능을 평가하였다. 그림 2-6은 SVM의 threshold 값에 따른 precision과 recall의 추이를 보이고 있다. Threshold 값을

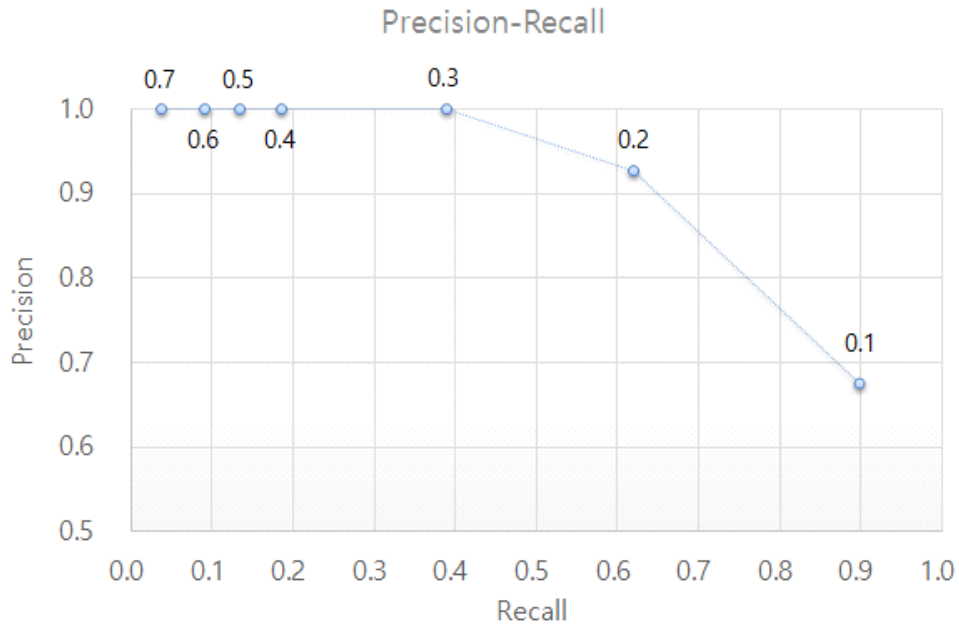


그림 2-6. Precision-Recall 추이

0.1에서 0.7까지 변경시키면서 총 7번의 테스트를 하였다. 일반적으로 threshold 값을 변경시키면 precision과 recall 값이 서로 trade-off 관계를 갖게 된다. 그림 2-6의 그래프에서도 이러한 현상을 관찰할 수 있다. 이 연구에서는 앞서 언급된 바와 같이 탐지 기능은 precision을 최대화하는 방향으로 설정하여야 하며, 그에 따라 threshold 값을 0.3으로 설정하는 것이 가장 적합한 것으로 나타났다. 0.3일 때, precision 100%를 확보하였고, 이보다 높은 threshold 값에서는 recall만 저하되는 것으로 나타났다. 그림 17은 선 자세의 템플릿과 쭈그린 자세의 템플릿을 동시에 사용하여 탐지한 결과의 예를 보이고 있다.

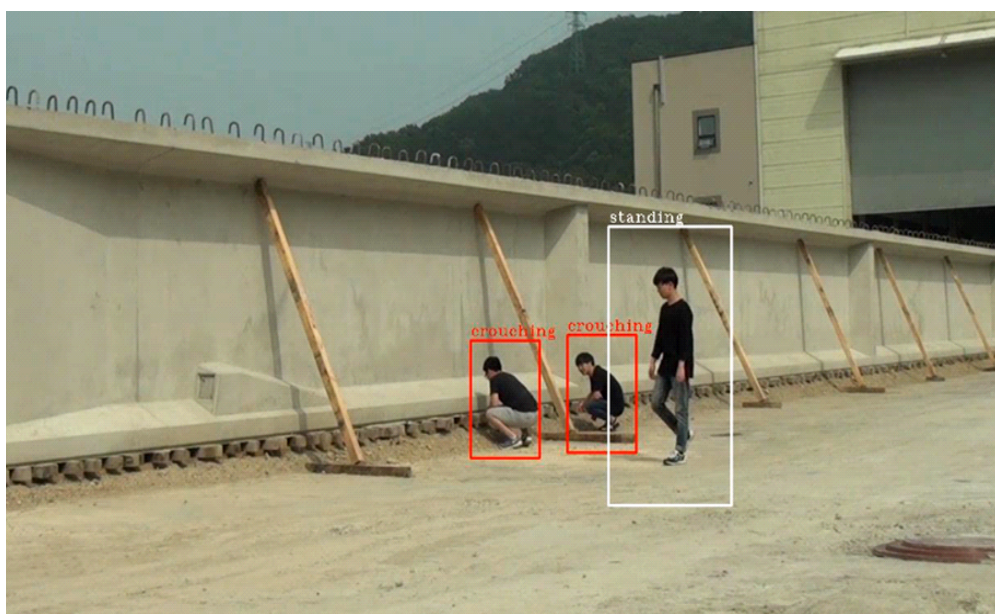


그림 2-7. 선 자세와 쭈그린 자세의 인력 탐지 예시

나. 안전모 탐지

현장 안전관리 차원에서 현장인력들의 안전모 착용 여부를 모니터링하기 위하여 안전모 탐지도 구현하였다. 안전조끼가 연녹색, 노란색, 주황색 등으로 색상이 비교적 한정되어 있는 것과 달리 안전모는 매우 다양한 색상이 존재하기 때문에 안전모 탐지는 형상 정보만을 사용하였다. 건설인력 탐지에서도 사용하였던 HOG와 SVM을 안전모 탐지에도 적용하였다. SVM의 훈련과정을 위해 500여 장의 안전모 사진을 수집하였다. 안전모의 HOG 템플릿 크기는 40(가로)×32(세로)로 정하였으며, 정면, 후면, 측면의 사진들을 모두 일괄적으로 훈련하여 하나의 템플릿을 생성하여 사용하였다. 건설인력 탐지에서와 마찬가지로 안전모 탐지의 성능 역시 precision과 recall로 평가하였다. 앞서 언급된 바와 같이 precision과 recall은 서로 trade-off 관계를 갖는다. 여기서는 precision을 최대화(recall 최소)한 결과와 recall을 최대화(precision 최소)한 결과 두 가지를 시도하여 건설인력의 안전모 착용여부 판단에 적용하였다.

표 2-1. 안전모 탐지 결과

	Precision 최대 시	Recall 최대 시
TP	2,247	2,984
FP	9	1,893
FN	838	100
Precision (%)	99.6	61.2
Recall (%)	72.8	96.8

- TP: 탐지된 안전모 개체 중 실제로 안전모인 개체 수
- FP: 탐지된 안전모 개체 중 실제로 안전모가 아닌 개체 수
- FN: 영상에 나타난 안전모 개체들 중 탐지되지 않은 개체 수

2. 안전모 및 안전조끼 미착용자 탐지

가. 건설인력 및 안전모 간의 매칭 방법

안전모 착용 여부를 판단하기 위해서 탐지된 건설인력 개체와 안전모 개체 간의 매칭 과정이 필요하다. 여기서는 안전조끼의 착용여부와 관계없이 모든 사람을 건설인력으로 고려하여 연구를 진행하였다. 그림 2-8의 좌측은 인력 탐지 결과와 안전모 탐지 결과의 한 예를 보이고 있다. 흰색과 검은색 사각형 영역은 각각 건설인력과 안전모의 위치와 크기를 나타낸다. 그림에서 탐지된 5명의 건설인력의 안전모 착용 여부를 알아내려면 그림 2-8의 우측과 같이 탐지된 안전모들이 어떤 건설인력의 안전모에 해당하는지 연결해주는 과정이 필요한 것이다. 그림에서 연결된 안전모가 없는 경우 안전모 미착용자로 구분되어 점선으로 영역이 표시된 것을 볼 수 있다. 이 절에서는 이러한 연결과정에 대한 내용을 상세하게 설명하고자 한다.



그림 2-8. 건설인력 및 안전모 탐지 결과

그림 2-8에서와 같이 건설인력 탐지 결과와 안전모 탐지 결과가 항상 실제 개체들의 위치와 정확하게 일치하다면 연결과정이 매우 간단할 수 있으나, 개체 탐지 방법들이 항상 완벽한 결과만을 주지는 않으며 때로는 실제 위치에서 다소 벗어난 영역을 탐지하거나 실제 크기보다 작은 영역만을 탐지하기도 한다. 이러한 모든 경우의 탐지 결과에 대하여 높은 정확도의 매칭 결과를 얻기 위하여 탐지된 건설인력 영역마다 안전모가 탐지될 수 있을 만한 후보영역을 설정하였다. 그림 2-9와 같이 두 가지 후보영역을 설정하였다. 좌측은 건설인력 탐지가 올바르게 이루어진 경우에 대한 설정이며 일반적으로 대부분의 경우가 좌측 설정에 의해 매칭이 이루어진다. Region I은 건설인력 탐지에 사용된 HOG 템플릿 영역의 구성에 맞추어 사람의 머리 부분이 위치할 수 있는 영역을 나타내고 있다. HOG 템플릿 영역은 그림 2-10과 같이 실제 인력의 위치 둘레로 좌우 $0.25w$, 상하 $0.125h$ 의 여백을 두고 있다. 이를 바탕으로 Region I의 영역을 설정하였다. 우측은 건설인력 탐지가 실제 개체 영역의 아래쪽으로 치우쳐진 경우에 대한 설정으로서 매우 드물게 발생

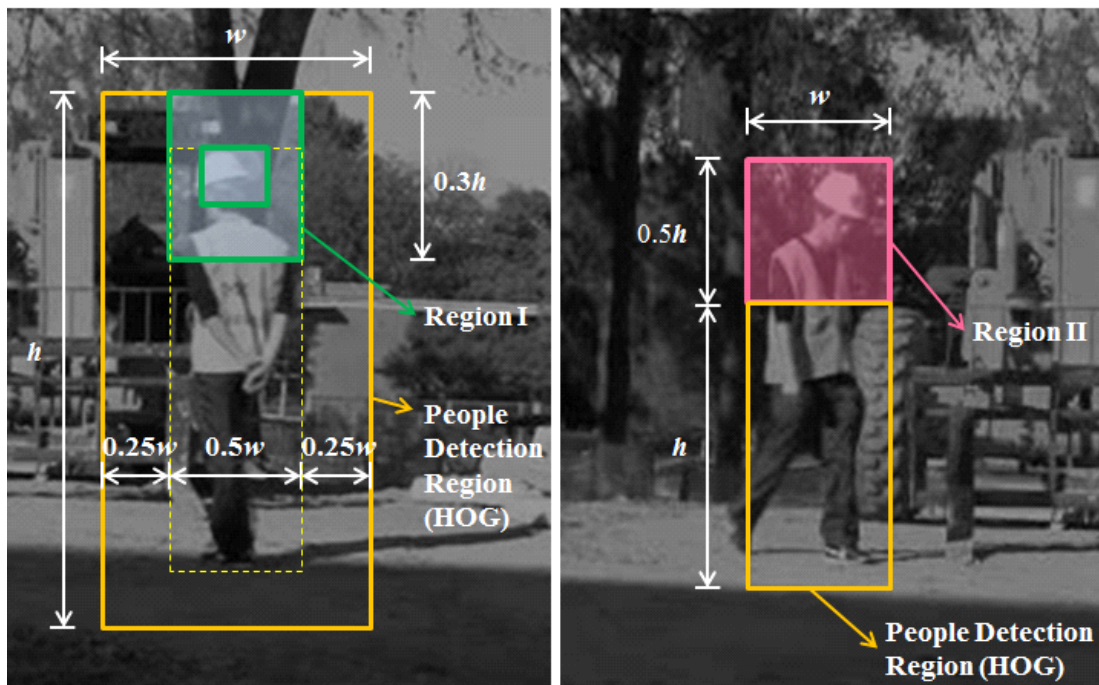


그림 2-9. 안전모 탐지 후보 영역 (좌측: Region I, 우측: Region II)

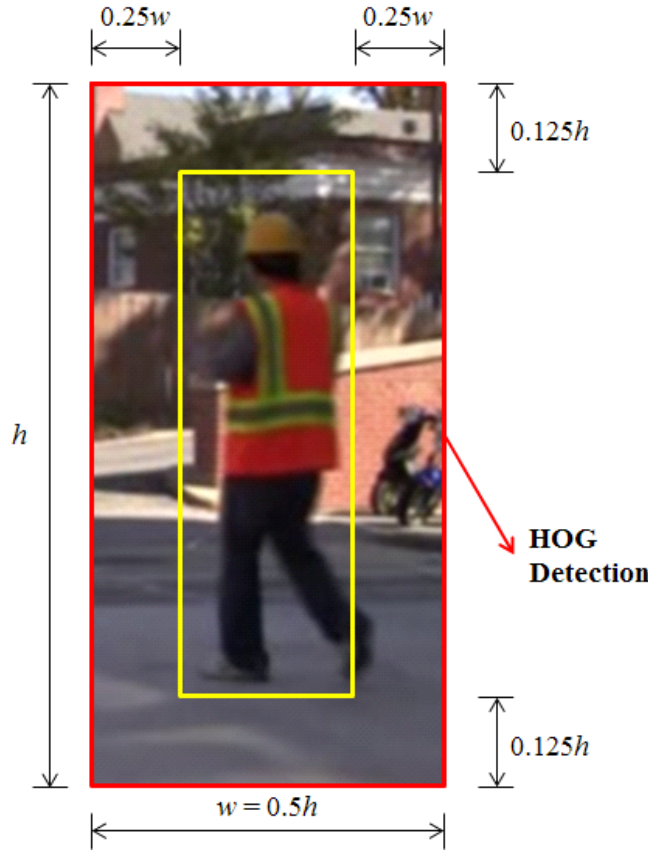


그림 2-10. 건설인력 탐지 HOG 템플릿 영역 구성

하는 경우이지만 정확도 향상을 위하여 추가된 방안이다. 실제로 건설인력 탐지의 결과를 조사해보았을 때, 건설인력이 카메라와 가깝게 위치해 화면에 비교적 크게 나타날 경우 그림 2-9의 우측과 같이 가슴 또는 허리 아래쪽 영역으로 탐지되는 경우가 존재하는 것을 확인하였기 때문에 우측과 같은 설정을 추가적으로 고안하였다. 알고리즘 자체적으로는 건설인력 탐지의 결과가 실제 개체 영역에 얼마나 정확하게 위치하는지 알 수 없기 때문에, 모든 탐지 결과에 대해서 Region I과 Region II를 고려하여야 한다.

건설인력 탐지 영역 P_i 의 Region I과 Region II에서 탐지된 모든 안전모 H_i 는 매칭이 가능한 후보 개체가 된다. 각 후보 개체들에 대하여 다음과 같이 정의된 D_{ij} 값을 계산하며 D_{ij} 가 최소인 안전모가 해당 건설인력의 안전모인 것으로 판단한다.

$$\text{Region I: } D_{ij} = a(1/s_{ij}) + b(dY_{ij})$$

$$\text{Region II: } D_{ij} = dY_{ij}$$

$$(s_{ij} = dy_{ij} / dx_{ij}, \quad dY_{ij} = dy_{ij} / w_j)$$

Region I의 경우 D_{ij} 는 s_{ij} 와 dY_{ij} 의 두 변수에 가중치 a 와 b 를 곱하여 합한 값으로 정의되었다. dx_{ij} , dy_{ij} , w_j 의 의미는 그림 2-11에 표현하였다. s_{ij} 와 dY_{ij} 는 두 점 p_1 과 p_2 간의 관계를 나타내는 변수로서 각각 두 점 간의 기울기와 정규화된 수직방향 거리를 나타낸다.

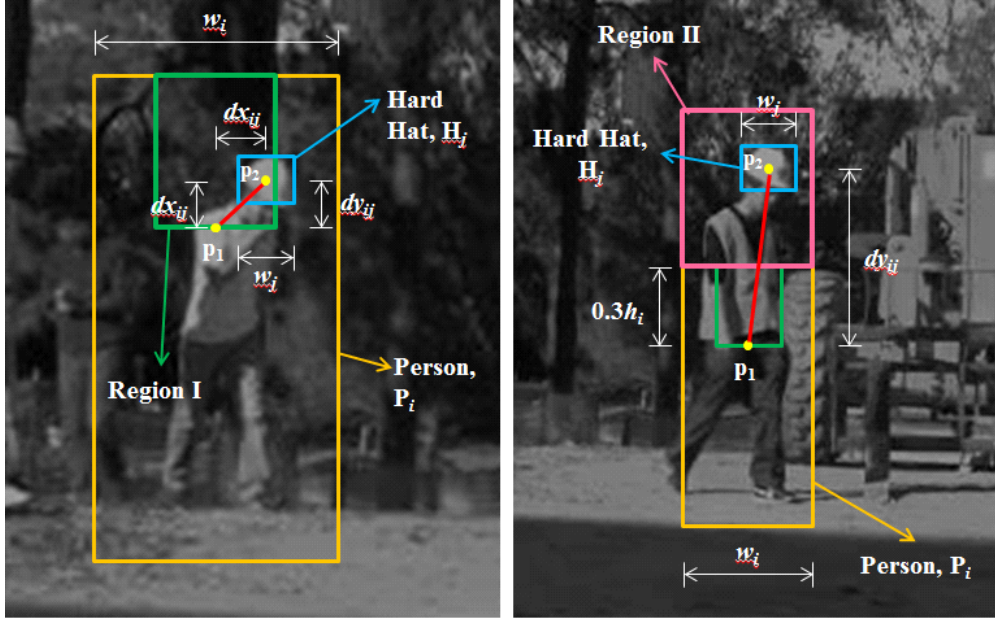


그림 2-11. 건설인력과 안전모 매칭을 위한 기준 변수
(좌측: Region I, 우측: Region II)

p_1 은 Region I의 하단 중심점이고 일반적으로 사람의 가슴 부분에 해당하며, P_2 는 탐지된 안전모의 영역의 중심점이다. 사람의 가슴 위치를 기준으로 안전모가 위치할 수 있을 만한 영역을 고려하여 변수 s_{ij} 와 dY_{ij} 가 D_{ij} 를 결정하는 요소로 선택되었다. 사람이 똑바로 있을 경우 $a(1/s_{ij})$ 는 0으로 수렴해 가고, $b(dY_{ij})$ 는 최대값을 갖게 된다. 사람이 고개를 숙이거나 꺾을 경우에는 반대로 $a(1/s_{ij})$ 는 증가하고, $b(dY_{ij})$ 는 감소하게 된다. 즉, 최소의 D_{ij} 를 매칭 안전모로 선택한다는 것은 고개를 똑바로 곧게 세우고 있어서 $a(1/s_{ij})$ 값이 최소이거나 고개를 숙여서 $a(1/s_{ij})$ 가 증가하는 만큼 $b(dY_{ij})$ 가 감소한 경우인 값을 찾는다는 것을 의미한다. s_{ij} 와 dY_{ij} 는 모두 무차원 값이기 때문에 건설인력 개체의 크기에 관계없이 동일한 매칭 성능을 기대할 수 있다. Region II에 대해서는 dY_{ij} 만 고려되었다. Region II에 나타난 안전모 영역의 $(1/s_{ij})$ 값 범위는 매우 좁고 한정적이기 때문에 안전모 위치에 따라 큰 차이가 나지 않으며 값의 크기 또한 dY_{ij} 에 비해 매우 작아 큰 영향을 주지 않는다.

최소의 D_{ij} 값을 선택하는 과정에 다음과 같은 추가적인 제약조건을 부여한다. 최소의 D_{ij} 를 갖는 안전모 개체 중에서도 다음의 조건을 만족하지 않는 경우 후보군에서 제외하여 매칭 오류를 줄이는 것이다.

$$\text{Region I: } (D_{ij}, dY_{ij}) \in A$$

$$\text{Region I: } (SR_{ij}, dY_{ij}) \in B, \quad (SR_{ij} = w_i / w_j)$$

영역 A와 B는 매칭 테스트 결과 데이터를 그림 2-12와 같이 분포도로 표시하고 오류 매칭 결과를 구분할 수 있는 경계선을 설정함으로써 결정하였다.

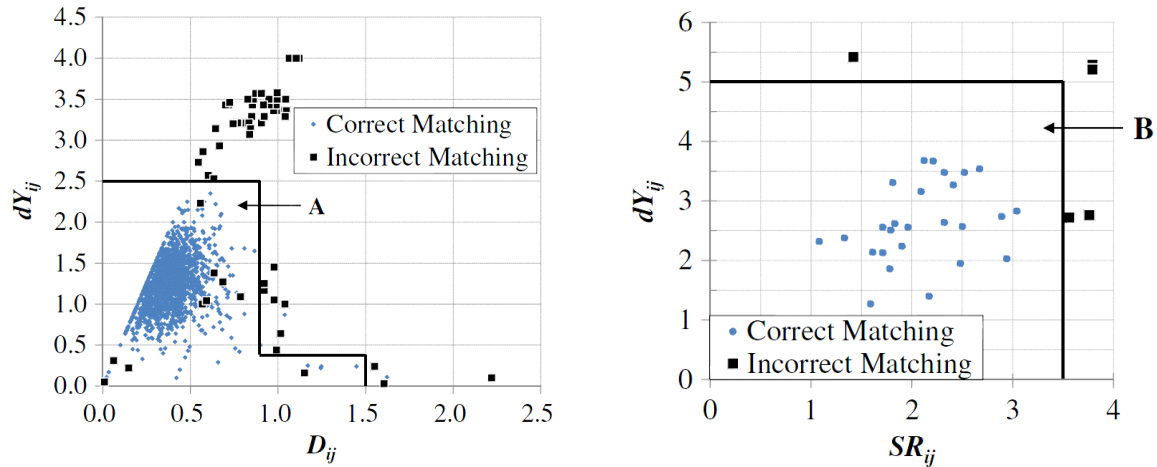


그림 2-12. 제약조건 영역 A와 B

나. 실험 결과

표 2-2. 건설인력과 안전모의 매칭

매칭 결과	안전모 탐지 알고리즘 설정	
	Precision 최대 시	Recall 최대 시
TP	2,542	2,027
FP	69	28
FN	1	1
Precision (%)	97.4	98.6
Recall (%)	100.0	99.8

- TP = 올바른 매칭 수
- FP = 잘못된 매칭 수 (다른 안전모와의 매칭)
- FN = 탐지된 건설인력 중 안전모를 착용하였음에도 안전모와 매칭이 형성되지 않은 경우의 수

안전모 및 안전조끼 착용여부에 대한 다양한 조합으로 인력을 구성하여 영상 데이터를 수집하였고, 이에 대한 매칭 테스트와 안전모 미착용자 탐지 성능 평가를 수행하였다. 테스트 영상은 길이 166초, 총 3,320 프레임으로 이루어졌고 해상도는 768×432이었다. 연산 시간은 총 332초가 소요되었다. 매칭 알고리즘의 성능에 대한 결과는 표 2-2에 정리되었다.

그림 2-13은 안전모 탐지가 최대의 recall을 갖도록 설정한 경우에 대한 매칭의 FP 및 FN 결과들을 보이고 있다. 첫 번째 그림은 안전모를 착용한 것으로 인식이 되긴 했으나 실제로는 뒤쪽에 위치한 다른 인력의 안전모와 매칭이 된 경우이다. 뒤쪽에 위치한 인력의 안전모 영역이 더 작은 D_{ij} 값을 갖으면서 제약조건 영역 A와 B 내에 들어온 경우이

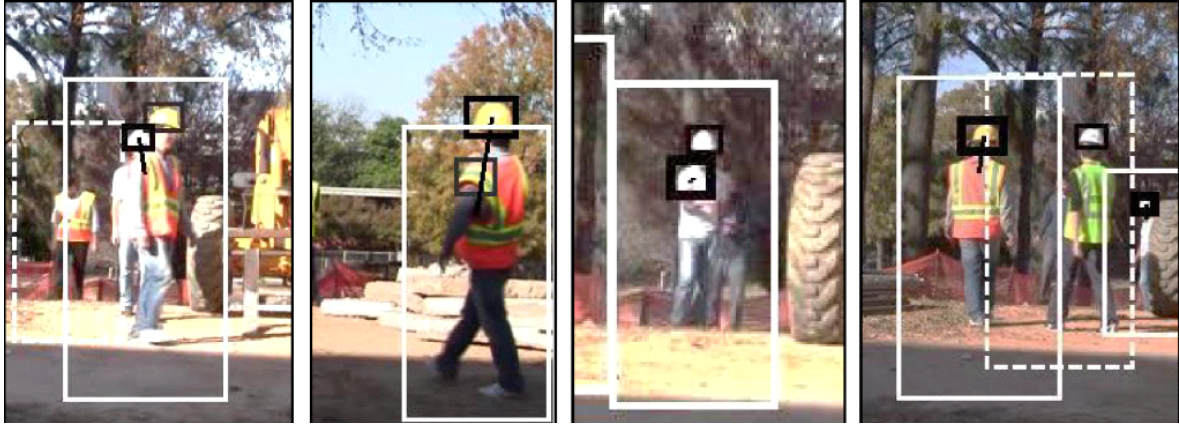


그림 2-13. 매칭의 FP 및 FN 결과들

다. 두 번째와 세 번째 그림들은 안전모 탐지의 FP 결과들에 의한 잘못된 매칭 결과이다. Recall을 최대로 설정하였기 때문에 그림과 같이 안전모의 FP 결과들이 다수 나타나게 되고, 이 경우 잘못된 매칭이 발생할 확률이 높아진다. 그림 속의 인력들이 안전모를 착용하고 있지 않아도 착용한 것으로 판단될 수 있었던 경우들이다. 네 번째 그림은 건설인력과 안전모가 모두 제대로 탐지되었음에도 매칭이 이루어지지 않은 경우를 보이고 있다. 이는 안전모 영역이 제약조건 A와 B 영역 밖에 위치한 경우로서 표 2-2의 유일한 FN이 이에 해당한다.

매칭 결과에 근거하여 안전모 미착용자를 탐지한 결과는 표 2-3과 같으며, 그림 2-14는 표본 결과 프레임들을 보이고 있다. 표에서 확인할 수 있듯이 안전모 탐지 알고리즘의 설



그림 2-14. 안전모 미착용자 탐지 결과

정에 따라 결과가 현저하게 차이가 나는 것을 볼 수 있다. 안전모 탐지의 recall이 높을 때 안전모 미착용자 탐지의 precision이 높아짐을 확인할 수 있다. 94.3% precision의 의미는 안전모 미착용자에게 경고 알람을 울릴 경우, 5.7%의 잘못된 경고 메시지가 울릴 수 있다는 것이며, 89.4% recall의 의미는 화면상에 나타난 미착용자 인력 중 10.6%의 인력을 간과하고 놓쳐서 경고를 울리지 못한다는 의미이다. 10.6%의 부분은 영상을 사용하면서 해결이 가능하다고 판단된다. 예를 들어, 첫 번째 프레임에서 탐지되지 않은 10.6%의 미착용자가 이어지는 후속 프레임들에서 탐지가 될 확률은 매우 높아 몇 초 이내에 경고 메시지를 보낼 수 있을 것이다.

표 2-3. 안전모 미착용자 탐지

안전모 미착용자 탐지 성능	안전모 탐지 알고리즘 설정	
	Precision 최대 시	Recall 최대 시
TP	932	921
FP	808	56
FN	131	109
Precision (%)	53.6	94.3
Recall (%)	87.7	89.4

- TP = 탐지된 안전모 미착용자 중 실제로 안전모 미착용자인 개체 수
- FP = 탐지된 안전모 미착용자 중 실제로 안전모 착용자인 개체 수
- FN = 안전모 미착용자 중 탐지되지 않은 개체 수

표 2-4. 탐지된 건설인력의 분류

인력	Type I	Type II	Type III	Type IV
안전조끼 착용	O	O	X	X
안전모 착용	O	X	O	X

표 2-5. TP, FP, FN, TN의 정의

Classified \ Actual	Type II, III, or IV	Type I
Type II, III, or IV	TP (1668)	FP (254)
Type I	FN (36)	TN (994)

마지막으로, 안전모 외에 안전조끼의 미착용 여부까지 고려하여 경고 알람을 울리기 위한 테스트를 수행하였다. 건설인력 탐지 알고리즘으로부터 안전조끼 착용자와 미착용자를 구분하여 결과를 얻고, 이를 앞서 소개한 안전모 미착용자 탐지 결과가 혼합하여 다음의 표 2-4와 같이 네 가지로 탐지된 인력을 분류하였다. Precision과 recall 계산을 위한 TP,

FP, FN, TN의 정의는 표 2-5와 같다.

안전모 탐지 알고리즘은 recall을 최대화한 경우를 사용하였으며, 표 2-4와 2-5에 근거하여 계산한 결과 precision과 recall은 각각 86.8%와 97.9%로 나타났다. 안전모나 안전조끼를 착용하지 않은 인력은 대부분(97.9%) 탐지가 되는 반면에, 경고 알람의 13.2%가 잘못된 알람이었다는 결과이다. 13.2%의 오류는 대부분 안전조끼 착용여부를 판단하는 분류과정에서의 오류에 기인한 것으로 분석되었으며, 잘못된 알람은 현장에서 혼란을 야기하고 작업효율을 오히려 저하시킬 수 있으므로 precision을 향상시킬 수 있는 방안이 요구된다.

3. Object Detection과 2D Tracking의 연계 알고리즘 개발

Object Detection은 기본적으로 각각의 프레임마다 화면상에 나타난 모든 건설인력들의 위치를 제공한다. 하지만, 프레임 간에 연계가 전혀 이루어지지 않은 프레임별 독립적 결과만을 제공할 뿐이다. 따라서, 각 인력 개체들 간의 구분이 이루어지지 않으며, 따라서 개체별 위치의 시간이력 데이터는 취할 수 없다. 개체별 추적, 즉 위치의 시간이력 데이터를 제공하기 위해서는 프레임과 프레임을 연관시켜주는 2D Tracking 과정이 필요하다(그림 2-15). 이 연구에서 제안하는 framework에서, object detection의 기본적 역할은 각 건설인력 개체가 처음 화면에 등장하였을 때, 인지하고 위치를 탐지하는 것이다. 한 번 탐지된 개체들은 이후에 2D tracking에 의해 위치 추적이 이루어진다. 하지만, 2D

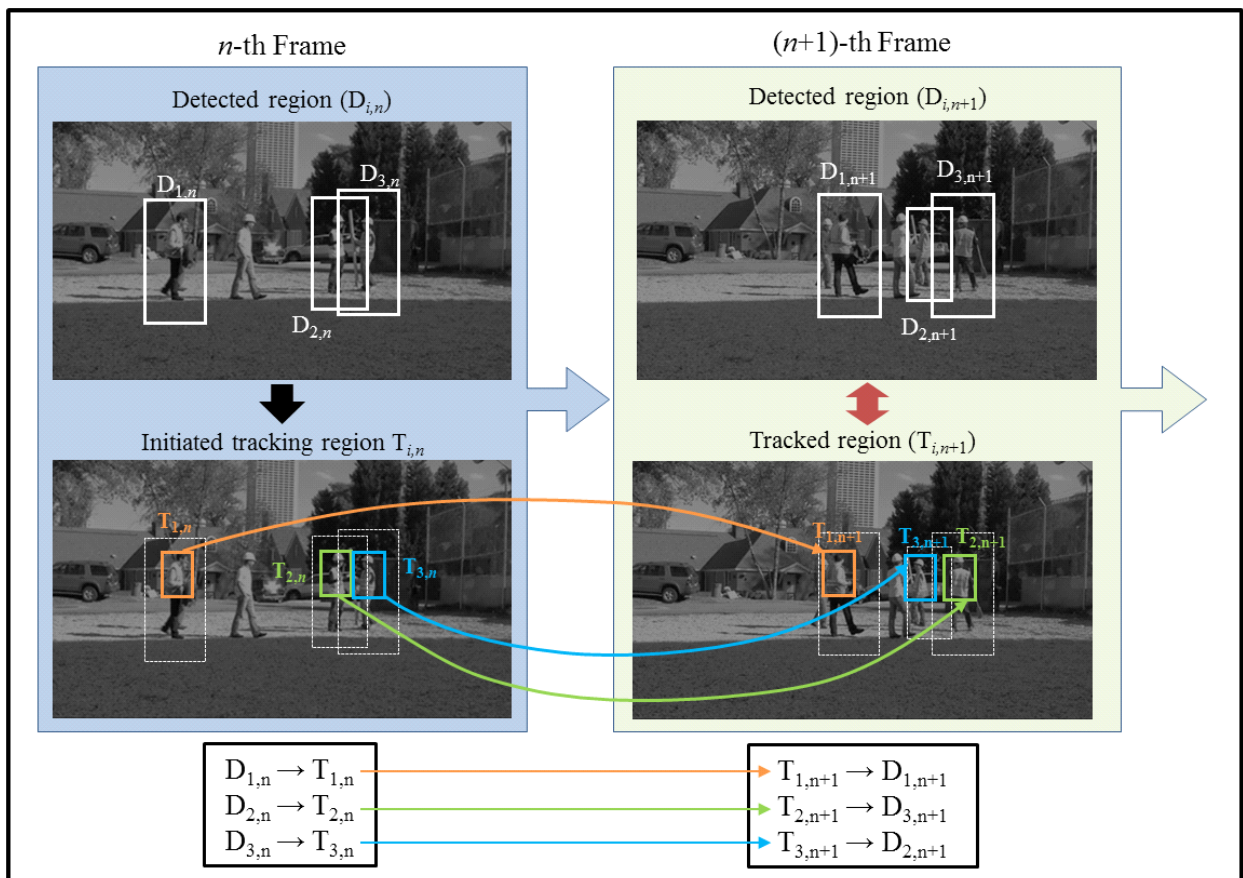


그림 2-15. 2D Tracking 알고리즘에 의한 프레임별 개체 탐지 결과들의 연계

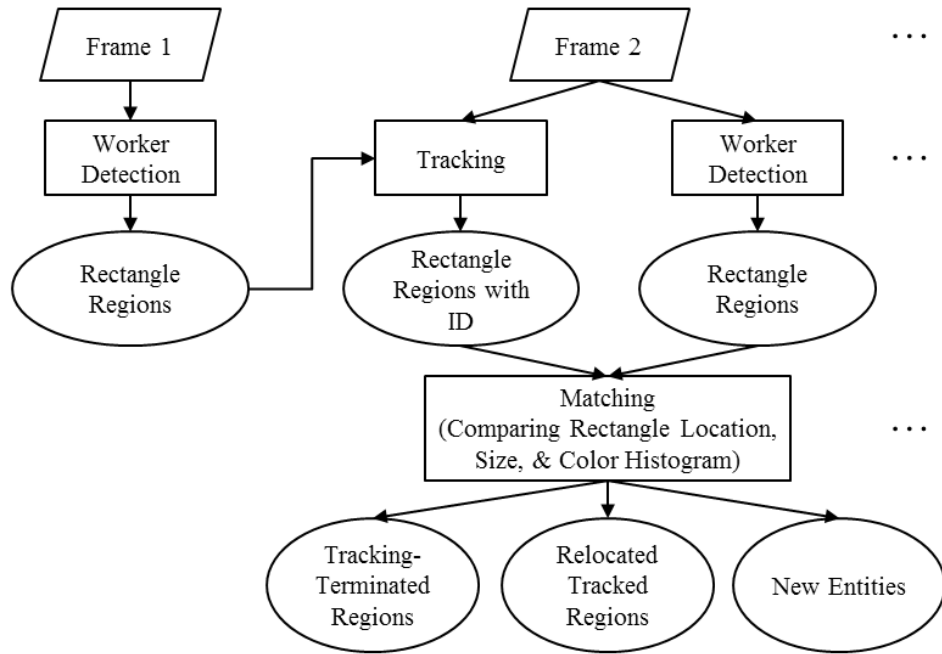


그림 2-16. 2D Tracking 알고리즘에 의한 프레임별 개체 탐지 결과들의 연계

tracking 알고리즘들은 추적 개체가 화면상에서 사라질 경우 오류를 일으키게 된다. 화면 밖으로 나가거나(exit) 다른 물체에 가릴 경우(occlusion), 또는 실내로 들어가는 경우가 이에 해당한다. 따라서 오류가 일어날 경우 미리 예측하여 2D tracking을 종료하고, 개체가 다시 화면상에 나타날 경우 object detection을 이용하여 탐지하고 2D tracking을 다시 개시해야 한다. 이러한 object detection과 2D tracking의 연동은 개체 수가 많을수록, 그리고 혼잡한 현장일수록 보다 높은 안정성을 요하게 된다. 이와 관련된 연구가 진행된 바가 있으나, 일부 건설 장비에만 한정된 연구였으며, 최대 2개의 개체만을 추적하는 결과만이 보고되었다. 또한, 2 개체마저도 서로 다른 종류의 장비(덤프트럭과 굴착기)이어서, 둘 사이의 구분이 명확한 시나리오에만 적용 가능하다. 이 연구에서는 같은 색의 안전모와 안전조끼를 착용하기 때문에 개체 간 구분이 어려운 건설인력을 다룬다.

그림 2-16은 건설인력 탐지 알고리즘과 2차원 위치추적 알고리즘의 연계를 위해 이 연구에서 제안하는 프레임워크이다. 첫 프레임에서는 건설인력 탐지 알고리즘만 적용된다. 이는 추적하고자 하는 개체를 찾기 위함이다. 건설인력이 탐지되면 각 탐지된 개체마다 독립적인 2차원 위치추적 프로세스가 시작된다. 건설인력 탐지 알고리즘은 모든 프레임에서 수행되는 반면에, 위치추적 알고리즘은 이전 프레임에서 추적된 개체가 존재하거나, 현재 프레임에서 새로 탐지된 개체가 존재할 때만 수행된다. 탐지 또는 추적된 개체는 개체를 포함하는 사각형 영역으로 표시되며, 각 사각형 영역마다 중심좌표, 크기, 내부영역의 color histogram의 데이터가 저장된다.

가. 건설인력 탐지 및 2차원 추적 알고리즘

건설인력 탐지 알고리즘은 앞 절에서 설명된 알고리즘을 사용하였으며 최종 3단계의 결과인 안전조끼 착용자에 대해서만 위치 추적을 수행하였다. 즉, 이 절에서의 “건설인력”은 안전조끼를 착용한 사람으로 정의한다. 그림 2-15 좌측 상단에서 건설인력 탐지 알고리즘의 결과는 건설인력 세 명의 영역 $D_{1,n}$, $D_{2,n}$, $D_{3,n}$ 이다. 마찬가지로, 다음 프레임에도 동일한 알고리즘이 적용되어 $D_{1,n+1}$, $D_{2,n+1}$, $D_{3,n+1}$ 의 결과를 얻는다. $D_{k,n}$ 에서 k 는 좌측으로부터 우측방향 순으로 주어진 번호이며 다른 의미는 없다. 이렇게 두 개의 프레임으로부터 얻은 dataset인 $D_{k,n}$ 과 $D_{k,n+1}$ 간의 연관 관계는 개체 탐지 알고리즘으로부터 얻을 수가 없다. 즉, 개체 탐지 알고리즘은 프레임 별로 독립적인 결과만을 생성하며, 시간이력 위치 데이터를 제공하지는 않는다. 앞 절에서 소개된 바와 같이 건설인력 탐지 알고리즘은 recall이 83.0%이므로 17%의 인력 개체를 놓치게 된다. 이러한 부족한 부분을 2차원 위치추적 알고리즘으로 대체하는 것이 이 연구의 목적이다.

2차원 위치 추적 알고리즘은 이전 프레임에서의 위치와 이미지 특성을 기반으로 현재 프레임에서의 위치를 예측하는 알고리즘을 말한다. 따라서 이전 프레임에서의 영역 정보만 입력 데이터로 주어져야만 시작이 가능하다. 한 번 시작이 되면 이후 이어지는 프레임에서는 매 프레임마다 결과를 제공한다. 개체가 화면에서 사라지더라도 추가적인 조치가 없으면 가장 비슷한 위치에서의 비슷한 외형 특성을 갖는 위치를 확률적으로 계산하여 결과로 제공한다. 이 연구에서는 책임연구자의 선행연구 성과인 논문인 “Comparative study of vision tracking methods for tracking of construction (Automation in Construction)”에 근거하여 템플릿 기반의 위치추적 알고리즘(Ross의 2008)을 채택/사용하였다. 사용된



그림 2-17. 2차원 위치 추적 영역의 설정

알고리즘은 eigen-image로 개체의 이미지 특성을 표현하고, particle filtering으로 새로운 영역의 위치 및 크기를 예측한다. 건설인력 탐지는 서 있는 사람의 형상을 이용하기 때문에 인력의 전신에 해당하는 영역을 결과로 준다. 반면에 위치 추적 시에는 그림 2-17에서 처럼 상반신 부분만을 추적 영역으로 설정하였다. 하반신 부분은 현장의 다른 물체들에 의해 가려질 확률이 높고, 다리의 움직임으로 인해 형상의 변화가 많기 때문이다. 그림 2-17에서는 HOG 템플릿 영역 중에서 여백 부분의 영역(좌우 $0.25W$, 상하 $0.125H$)을 제외한 후, 위쪽 절반만을 위치 추적 영역으로 설정하고 있음을 볼 수 있다.

나. 두 알고리즘의 연계

개체 탐지 알고리즘과 2차원 추적 알고리즘은 각각 독립적인 결과 데이터를 생성한다. 두 데이터 모두 건설인력 개체의 영역을 나타내는 사각형 영역으로 표현된다. 이 두 데이터 그룹을 통합하기 위해서 간단한 비교 과정을 거친다. 각각의 2차원 위치 추적 영역마다 모든 개체 탐지 영역과의 비교를 통하여 사각형 영역이 서로 겹치는지의 여부를 보고 동일 개체를 나타내는 사각형 영역인지를 판단한다. 겹치는 영역이 두 개 이상일 경우, 영역 중심점 간의 거리를 계산하여 가장 가까운 영역을 선택한다. 이러한 비교 과정의 결과는 다음과 같은 세 가지 경우로 나누어 볼 수 있다.

- (1) 위치추적 영역과 개체탐지 영역이 짝을 맺은 경우
- (2) 매칭되는 개체탐지 영역이 없는 위치추적 영역
- (3) 매칭되는 위치추적 영역이 없는 개체탐지 영역

첫 번째 경우에는 개체탐지 영역을 위치추적의 새로운 시작 영역으로 재설정한다. 일반적으로 위치추적의 결과는 개체탐지에 비해 불안정한 면이 있으며, 개체탐지 알고리즘은 precision이 99.0% 이상이 되도록 설정하였으므로 그 결과의 신뢰도가 높기 때문이다. 그

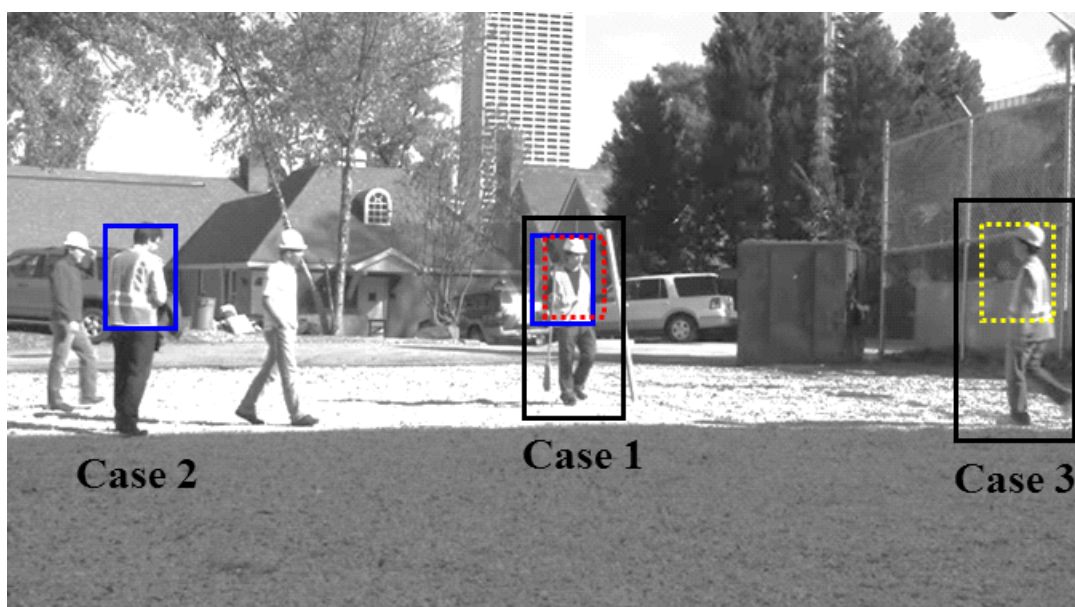


그림 2-18 개체 탐지와 위치추적 결과의 비교 과정의 세 가지 경우

림 2-15에서는 세 개체가 모두 첫 번째 경우에 해당하며 이 때 2차원 위치추적 알고리즘의 주된 역할은 연속된 프레임으로부터 얻은 개체탐지 결과를 연결시켜주는 것이다.

두 번째 경우에서의 2차원 위치추적 알고리즘의 주된 기능은 개체탐지 알고리즘이 탐지하지 못하고 놓친 개체들을 연속적으로 추적하는 것이다. 이는 81%에 그치는 개체탐지 알고리즘의 recall 값에 대한 보완 효과를 만들어낸다.

세 번째와 같이 개체탐지 영역이 단독으로 나타난 경우는 이전 프레임에서는 없었던 새로운 개체가 나타났음을 의미한다. 새롭게 탐지된 개체는 새로운 위치추적 프로세스의 개시로 이어진다. 예를 들어, 세 개의 개체가 새로 탐지된다면, 세 개의 새로운 위치 추적 프로세스가 추가적으로 시작된다. 그림 2-18은 설명된 세 가지 경우의 예를 하나씩 보이고 있다.

앞서 언급된 바와 같이 위치추적 알고리즘은 매 프레임마다 결과를 생성한다. 이 위치추적 알고리즘은 개체가 카메라 영상 화명 밖으로 나갈 경우 종료하여야 한다. 강제로 종료하지 않으면 불안정한 상태로 잘못된 영역 결과를 지속적으로 얻게 된다. 따라서 추적 영역이 영상 프레임의 가장자리에 닿을 경우 해당 개체의 위치추적 프로세스를 자동 종료하도록 설정한다. 이는 다음과 같이 표현할 수 있다.

$$x_{min} < 0 \text{ 또는 } x_{max} \geq W \text{ 또는 } y_{min} < 0 \text{ 또는 } y_{max} \geq H \text{ 이면}$$

위치추적 프로세스 종료

x_{min} 과 x_{max} 는 각각 영역의 최소 및 최대 x 좌표를 나타내며, y_{min} 과 y_{max} 는 각각 영역의 최소 및 최대 y 좌표를 나타낸다. W 와 H 는 영상 프레임의 가로 및 세로 크기이다. 일반적으로 프레임의 좌측 상단이 이미지 좌표의 원점이며, 우측 하단의 좌표는 $(W-1, H-1)$ 이 된다.

사용된 건설인력 탐지 알고리즘의 precision이 99% 정도가 되도록 설정되어 탐지된 개체들은 거의 대부분 실제로도 건설인력이지만, 여전히 1%의 확률로 건설인력이 아닌 개체가 탐지되어 추적될 가능성이 있다. 한 번 탐지된 개체는 위치추적 알고리즘에 의해 후속 프레임에서도 연속적으로 결과를 생성하게 된다. 즉, FP의 수가 계속해서 축적되게 된다. 이러한 문제점을 해결하기 위하여 FP 개체임을 감지하고 추적 프로세스를 자동 종료할 수 있는 기능을 고안하였다. 1%의 확률로 탐지된 개체가 후속 프레임에서 지속적으로 다시 탐지될 확률은 매우 낮다. 따라서 “일정 시간” 이상 매칭되는 개체 탐지 영역이 없는 위치추적 프로세스는 FP 개체로 판단하고 위치추적 알고리즘을 종료한다. 여기서 중요한 부분은 종료의 기준이 되는 “일정 시간”의 길이를 결정하는 것이다. 이 길이를 길게 설정하면 FP 개체가 추적되는 시간이 그만큼 길어지며, 반대로 짧게 설정하면 TP 개체인 건설인력 개체의 추적 프로세스까지도 종료되는 경우가 생길 수 있다. 건설인력 개체들의 평균적인 탐지 빈도수를 분석하여 최종적으로 3초를 기준으로 결정하였다.

다. Occlusion 처리 방법

복잡한 환경의 건설현장에서는 건설인력이 다른 물체에 가려져 영상의 시야에서 사라지는 경우가 자주 발생할 수 있다. 화면에서 사라진 개체를 지속적이고 안정적으로 추적할 수 있는 방법은 없다. 가려진 상태에서의 위치나 외형의 변화를 예측하기가 어렵고 신뢰도가 낮기 때문이다. 추적하던 개체가 가려져서 안 보이게 될 경우, 위치추적 알고리즘은 매우 불안정한 상태가 되며 잘못된 결과를 제공하게 된다. 따라서 이 연구에서는 이렇게 가려져서 안 보이게 되는 occlusion이 발생할 경우, 이를 최대한 빨리 감지하여서 추적 프로세스를 종료하는 방법을 고안하였다. Occlusion은 두 가지 경우로 나뉘볼 수 있다. 첫 번째(OC1)는 건설인력이 아닌 다른 물체에 의해 가려지는 경우이다. 예를 들어 건설 장비나 현장에 적재되어 있는 자재들에 의해 가려지는 경우가 있다. 그림 2-19가 이러한 예를 보이고 있다. (n-2) 번째 프레임에서 추적되고 있는 건설인력 개체가 (n-1) 번째 프레임에서부터 일부가 가려지기 시작하고 있다. 이 건설인력 개체가 장비 뒤쪽으로 사라져 보이지 않는 상황에서 별도의 종료 프로세스를 두지 않으면, 이 개체의 추적 영역은 n 번째 프레임 이후로는 그림의 점선으로 표시된 영역에 계속 머무르게 된다. 이러한 OC1 상태는 다음의 두 가지 요소를 바탕으로 감지한다.

- (1) 매칭되는 개체탐지 영역이 없음
- (2) color histogram이 급격하게 변함

건설인력 탐지 알고리즘은 전신의 일부만 가려져도 탐지 확률이 매우 낮아진다. 따라서 그림 2-19의 (n-1) 번째와 n 번째 프레임에서는 개체 탐지 결과를 얻지 못하게 되며, n 번째 프레임에서는 영역의 대부분이 장비의 바퀴 부분으로 채워져 색상이 급격하게 변하게 된다. 이러한 색상의 변화는 color histogram을 사용하여 감지한다. 각 위치추적 영역의 color histogram 매 프레임마다 다음과 같이 계산된다.

$$C_n = (C_{n-1} + c_n) / 2$$

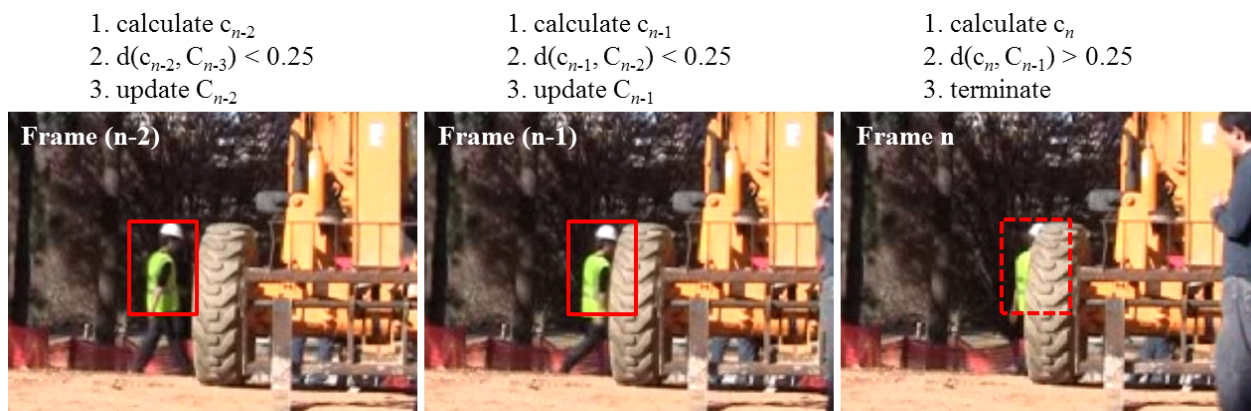


그림 2-19. 건설 장비에 의해 발생한 occlusion을 인식하는 과정

여기서 c_n 은 n 번째 프레임에서의 위치추적 영역에 대한 color histogram이고, C_n 은 color histogram의 가중 평균이다. 가중 평균의 가중치는 현재 영역의 color histogram에 0.5를 부여하고 이전까지 누적된 가중 평균에 0.5를 부여하여 현재의 색상에 비중을 많이 두고 있다. 이렇게 매 프레임 C_n 을 계산하기 이전에 C_{n-1} 과 c_n 의 Euclidean 거리를 계산하여 0.25보다 큰 경우를 occlusion 발생으로 판단한다. 기준값이 되는 0.25는 반복적인 테스트를 통하여 산정하였다.

Occlusion의 두 번째 종류(OC2)는 추적되고 있는 두 건설인력 개체끼리 중첩되어 이 중 하나의 개체가 가려지는 경우이다. 그림 2-20이 한 예를 보이고 있다. OC2 상황을 처리하기 위하여 서로 중첩된 정도를 다음과 같이 세 가지 단계로 구분하였다.

Level 1: $0 < p \leq 0.2$

Level 2: $0.2 < p \leq 0.6$

Level 3: $p > 0.6$

여기서 p 는 두 영역 중 크기가 작은 영역의 면적 대비 중첩된 부분의 면적의 비율이다. 그림 2-20의 세 그림은 좌측부터 각각 이 세 단계를 나타내고 있다. Level 1 상태는 두 개체 간의 구분이 비교적 명확한 상태이므로 occlusion으로 분류하지 않는다. 따라서 일반적인 프로세스가 그대로 진행된다. Level 2에서는 그림에서 보듯이 두 개체가 하나의 개체 탐지 결과를 공유하게 되며 어느 쪽이 올바른 매칭인지 판단하기가 어려운 상태이다. 그렇기 때문에 Level 2 상태에서는 개체 탐지 결과와의 매칭 과정을 생략한다. 사용된 위치추적 알고리즘은 60% 정도가 가려진 상태에서도 안정적으로 위치추적이 가능한 것을 확인하였기 때문에 개체 탐지 결과와의 매칭이 없어도 문제가 없는 것으로 확인되었다. Level 3 상태가 되면 두 개체 중 하나는 거의 보이지 않게 되며 위치추적 결과도 불안정한 상태로 전환되기 시작한다. 이 상태에서는 두 개체 중 어느 개체가 가려진 상태인지를 판단하여 해당 개체의 위치추적 프로세스를 종료하여야 한다. 그림 2-20(c)에서는 Tracking 1이 Detection 영역과 매칭되어야 하며, Tracking 2의 프로세스는 종료되어야 한다. 이 시점에서의 매칭은 다음의 세 가지 거리값에 의하여 결정된다.

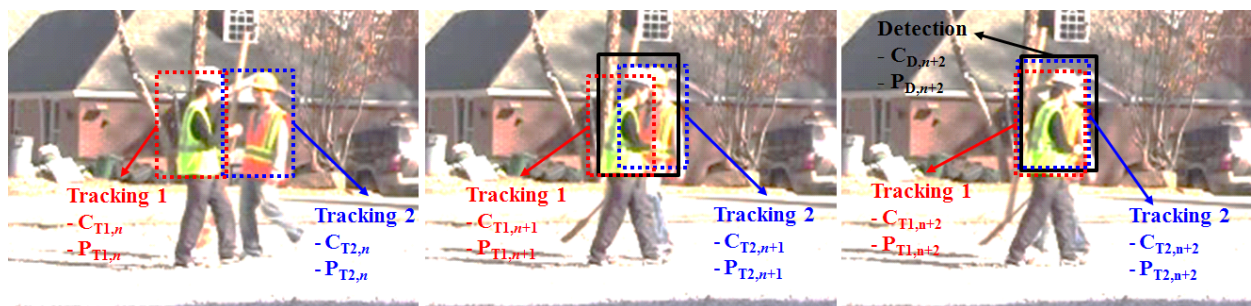


그림 2-20. 건설 장비에 의해 발생한 occlusion을 인식하는 과정

$$\begin{aligned}
d_{P,1} &= |\mathbf{P}_{D,n+1} - \mathbf{P}_{T1,n+1}|, & d_{P,2} &= |\mathbf{P}_{D,n+1} - \mathbf{P}_{T2,n+1}| \\
d_{C_dt,1} &= |\mathbf{C}_{D,n+1} - \mathbf{C}_{T1,n}|, & d_{C_dt,2} &= |\mathbf{C}_{D,n+1} - \mathbf{C}_{T2,n}| \\
d_{C_tt,1} &= |\mathbf{c}_{T1,n+1} - \mathbf{C}_{T1,n}|, & d_{C_tt,2} &= |\mathbf{c}_{T2,n+1} - \mathbf{C}_{T2,n}|
\end{aligned}$$

여기서 $\mathbf{P}_{D,i}$ 는 개체 탐지 영역의 (x, y, w) 이고, $\mathbf{P}_{Tj,i}$ 는 j 번째 위치추적 영역의 (x, y, w) 이다. x, y, w 는 중심좌표와 영역의 폭을 나타낸다. 영역의 거리 외에도 크기도 중요한 비교 속성으로 고려하기 위하여 w 가 포함되었다. 마찬가지로, $\mathbf{C}_{D,i}$ 와 $\mathbf{c}_{Tj,i}$ 는 i 번째 프레임에서의 개체탐지 영역과 위치추적 영역의 color histogram이고, $\mathbf{C}_{Tj,i}$ 는 가중 평균 i 번째 프레임에서의 가중평균 color histogram이다. 첫 번째와 두 번째 식은 개체탐지 영역과 위치추적 영역 간의 공간적 거리와 색상의 차이를 구하는 것인 반면에, 세 번째 식은 위치추적 영역 자체의 색상 변화를 계산하고 있다. 그림 2-20에서 파란색 영역인 Tracking 2는 $(n-1)$ 번째 프레임에서 n 번째 프레임으로 넘어가는 과정에서 주황색에서 연녹색 위주의 색으로 급격한 색상 변화가 생겼다. 이러한 색상의 급격한 변화는 occlusion의 단서가 되며 세 번째 식에 반영된다. 위의 식에 의해 계산된 거리 $d_{P,i}$, $d_{C_dt,i}$, $d_{C_tt,i}$ 중 더 큰 값이 두 개 이상인 개체가 가려진 개체로 결정되어 위치추적 프로세스가 종료되게 된다.

라. 실험 결과

제안된 프레임워크를 Microsoft Visual C#으로 프로그래밍하여 테스트 영상들에 적용하여 성능을 평가하였다. 위치추적 알고리즘과의 연계를 통해 성능이 향상되었음을 보이기 위해 개체탐지 알고리즘만을 사용하였을 때와의 결과와 비교해 보았다. 성능 평가는 precision, recall, accuracy를 기준으로 하였다. Accuracy는 다음과 같이 정의된다. 비교 결과는 표 2-6에 정리하였다.

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN})$$

- TP = 탐지/추적된 건설인력 중 실제로 건설인력인 개체 수
- FP = 탐지/추적된 건설인력 중 실제로 건설인력이 아닌 개체 수
- FN = 건설인력 중 탐지/추적되지 않은 개체 수
- TN = 안전조끼를 착용하지 않았기 때문에 탐지/추적되지 않은 개체 수

표 2-6에 나타난 바와 같이 precision이 0.8% 저하되었지만 recall이 15.0% 만큼 향상되었고, 결과적으로 accuracy도 7.7% 향상된 결과를 얻을 수 있었다. Precision이 감소된 원인은 건설개체가 아닌 개체(안전조끼 미착용자)가 탐지되어 1~3초 정도 추적되었기 때문이다. Recall이 98.0%라는 것은 2% 만큼 놓친 개체가 있음을 의미하는데 이는 대부분 새로운 개체가 나타난 후 바로 탐지되지 않고 몇 프레임 이후에 탐지되는 경우가 있기 때문이었다. 그림 2-21은 결과 프레임들의 몇 가지 표본들이다. 추적된 건설인력은 흰색 사각

형으로 표시되었으며 각각의 영역은 추적이 시작된 순서대로 번호 레이블이 부여되었다. 그림 2-21 상단의 두 그림은 각각 FP와 FN의 결과를 하나씩 포함하고 있다. 상단 좌측의 그림에는 안전조끼를 착용하지 않은 개체가 추적되고 있음을 볼 수 있고(FP), 우측 상단의 그림에는 안전조끼를 착용하였음에서 추적되지 않은 개체를 볼 수 있다(FN).

표 2-6. 프레임 단위로 산정한 precision, recall, accuracy

	Precision (%)	Recall (%)	Accuracy (%)
개체 탐지 + 2D 위치 추적	98.9	98.0	98.2
개체 탐지	99.7	83.0	90.5
증감	-0.8	15.0	7.7

제안된 방법에서는 이전 프레임에서 추적되었던 개체가 화면에서 사라졌다가 다시 나타날 경우 새로운 개체로 인식한다. 일반적으로 사라지기 전의 개체와 다시 나타난 개체를



그림 2-21. 건설인력 탐지 알고리즘과 2차원 위치추적 알고리즘의 연계에 의한 건설인력 위치 추적 결과 프레임의 표본
동일한 개체로 인식하기에는 정보가 부족하여 매칭이 어렵다. 따라서 새롭게 탐지된 개체를 모두 새로운 건설인력으로 간주하였고 총 89개의 개체가 탐지/추적되었다. 이 중 82개



그림 2-22. 안전조끼 미착용자인 15번 개체(FP)의 탐지(좌측 그림), FP의 추적 과정 (중간 그림), FP 추적의 자동 종료 (우측 그림)

의 개체가 실제 건설인력이었으며, 7개의 개체는 사람이 아니거나 안전조끼 미착용자였다. 영상에 나타난 건설인력이 단 하나의 프레임에서도 탐지/추적되지 않은 경우는 없었다. 따라서 개체 단위로 precision과 recall을 계산하면 표 2-7과 같다.

표 2-7. 건설인력 개체 단위로 산정한 precision, recall

방법	Precision (%)	Recall (%)
개체 탐지 + 2D 위치 추적	82/89 = 92.1	82/82 = 100.0

테스트 영상에서 추적되던 개체가 화면 밖으로 나가는 경우는 24번 발견되었다. 제안된 방법은 26번 모두를 탐지하여 위치추적 프로세스를 자동 종료하였다. 또한 건설인력이 아닌 개체가 탐지되어 추적된 경우가 7번 발생하였으며, 이들의 위치추적 프로세스는 평균 1.9초 이내에 자동 종료되었다. 자동 종료까지 걸린 최대 시간은 3.8초였다. 그림 2-22는 안전조끼를 착용하지 않은 사람이 탐지/추적되었으나 1.7초 만에 종료되고 있음을 보이고 있다. 이러한 자동 종료 기능에 의해 지속적으로 추적되어야 할 건설인력의 위치추적 프로세스가 종료되는 경우도 세 번 발생하였다. 하지만 이 중 두 번은 1초 이내에 다시 탐지되었고, 이전에 추적되었던 개체와의 유사성이 발견되어 같은 개체로 인식되었고 같은 번호 ID를 부여받았다. 그림 2-23에서 16번 개체의 위치추적이 두 번째 그림에서는 중단되었다가 세 번째 그림에서 다시 탐지되어 이전과 같은 16번 ID를 부여받았음을 확인할 수 있다.

Occlusion은 OC1인 경우가 32번, OC2인 경우가 20번 발생하였으며 OC2에 해당하는 한 번의 케이스를 제외한 모든 경우가 안정적으로 처리되었다. 즉, 가려진 개체들을 즉각적으로 감지하고 위치추적 프로세스를 종료하였고, 다시 나타날 경우 새롭게 탐지하여 위치추적 프로세스를 시작하였다. 다만 20번의 OC2 케이스 중 한 번의 경우에 대해서 오류가 발생하였는데, 뒤쪽에 가려진 개체가 아니라 앞쪽에 위치한 개체가 가려진 것으로 판단되어 위치추적 프로세스가 종료되었다. 그림 2-24는 occlusion 처리 결과의 몇 가지 예를



그림 2-23. 16번 개체의 추적 (좌측 그림), 추적과정의 잘못된 자동 종료 (중간 그림), 다시 탐지되어 같은 번호를 부여받은 16번 개체 (우측 그림)

보이고 있다.

앞서 언급된 바 있듯이 한 개체가 사라졌다가 다시 나타날 경우 일반적으로 새로운 개체로 인식한다. 하지만, 매우 제한된 조건 하에서는 이전에 추적이 종료된 개체와 새롭게 탐지된 개체 간의 매칭이 가능하기도 하다. 사라지기 직전과 다시 나타났을 때의 시간 간격이 짧고 위치와 색상 정보가 매우 유사한 경우에 한하여 같은 개체로 인식할 수 있는 매칭이 가능하였다. 그림 2-23과 그림 2-24(d)의 경우가 이에 해당한다. 그림 2-24(d)에서 0번 개체가 사라졌다가 다시 매우 유사한 위치에 나타났고 색상의 변화도 미미하여 같은 개체로 인식되었음을 볼 수 있다. 종료된 위치추적 프로세스의 정보를 데이터베이스에 종료 후 일정시간 동안만 보관하여 새롭게 탐지된 개체 정보와 비교과정을 거쳐 매칭을 찾게 된다. 매칭의 기준 속성은 영역의 중심좌표, 크기, color histogram이었다. 동일한 개체



그림 2-24. Occlusion 처리 과정: OC1 (a, b), OC2 (c, d)



그림 2-26. Entity matching의 예

가 사라졌다가 다시 나타난 케이스가 총 39번 발견되었으며 이 중 언급된 제한된 조건에 해당하여 실제로 매칭이 이뤄진 케이스는 10번이었다.

4. Epipolar geometry 기반의 Entity matching

가. Entity matching의 목적 및 필요성

3차원 위치 좌표를 구하기 위해서는 두 개의 카메라가 필요하다. 개체와 두 카메라를 이용하여, 삼각 측량을 통해 개체의 3차원 좌표를 추정할 수 있다. 이에 앞서, 각 영상에서 추적하고자 하는 개체를 검출하고, 두 카메라의 기하학적 상관관계를 이용하여 각 영상에서 검출된 개체가 동일한 개체인지를 확인해야 한다. 이러한 과정을 개체 매칭(Entity matching)이라 한다. 이 절에서는 각 영상으로부터 동일한 개체를 검출하기 위해 사용 가능한 방법들에 대해 소개하고, 적절한 개체 매칭 방법을 제시한다. 최종적으로 개체 매칭의 정확도를 측정하여 실제 건설 현장에서의 개체 매칭 적용 가능성을 확인한다.

영상 기반 위치 추적 시스템은 기본적으로 두 개의 고정된 카메라를 필요로 한다. 또한, 추적하고자 하는 개체가 두 개의 카메라에 모두 포착될 수 있어야 하므로 카메라는 추적하고자 하는 개체를 포함하는 공통된 지역을 촬영한다. 그림 2-25에 영상 기반 3차원 위치 추적 시스템의 수행 방법을 간단하게 표현했다. 촬영된 목표 추적 개체의 2차원 영상

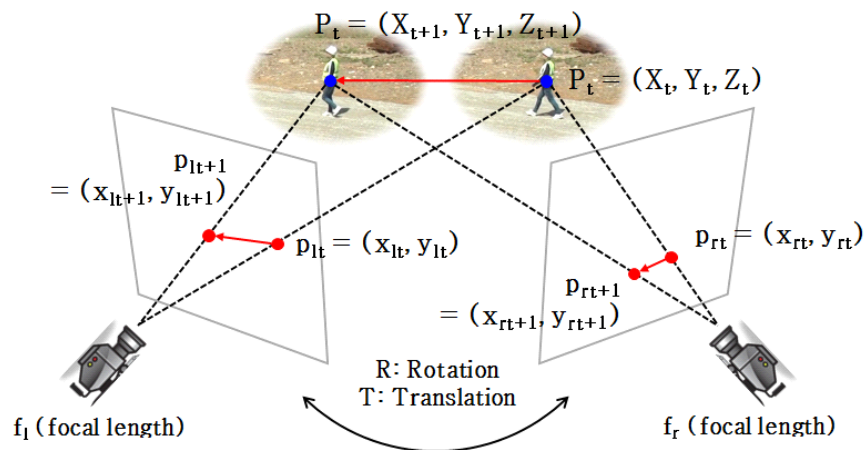


그림 2-25. 영상 기반 3차원 위치 추적의 개념도

좌표와 고정된 카메라의 기하학적 상관관계를 이용하면 삼각 측량을 통해 목표 추적 개체의 실제 3차원 위치 좌표를 계산할 수 있다.

추적하고자 하는 개체가 영상에 등장하는 유일한 개체이면 문제가 없지만 이런 경우는 극히 드물며 일반적으로 다수의 개체들을 추적하게 된다. 따라서 양쪽 카메라에서 탐지된 다수의 개체들 간에 매칭 작업이 필요하다. 즉, 좌측 영상에서 탐지된 특정 개체가 우측 영상에서 탐지된 개체들 중 어느 개체에 해당하는지를 알아내어야 한다.

본 연구는 개체 매칭 대상을 건설 인력으로 한다. 그림 2-26은 건설 인력 대상 개체 매칭의 예시이다. 좌측 영상과 우측 영상은 공통된 지역을 서로 다른 위치에서 촬영한 것이다. 우측 영상의 후보 개체들 중에서 좌측 영상의 목표 개체와 동일한 개체를 찾아내는 것이 개체 매칭이다. 개체 매칭은 카메라와 대상 간의 삼각 측량을 통해 3차원 좌표를 계산하기 위해서 반드시 필요한 과정이다. 두 개의 카메라를 이용하여 촬영된 영상은 각기 다른 화각과 축척을 가지기 때문에 개체 매칭을 위해서는 별도의 처리 과정이 필요하다. 이 연구에서는 그 첫 단계로서 등극선 기하학(Epipolar geometry)을 이용한 방법을 구현하고자 한다. 등극선 기하학은 두 카메라의 기하학적 상관관계를 나타내는 기초행렬(Fundamental matrix)로 표현 가능하며 영상에서 추출된 특징점들을 사용하면 보다 정확한 기초행렬을 계산할 수 있다.

나. Epipolar geometry 기반 entity matching

(1) 정의 및 적용 방법

그림 2-27은 이 연구에서 제안하는 등극선 기하학 기반 개체 매칭의 전반적인 수행 과정을 나타낸다. 등극선 기하학 기반의 개체 매칭은 기초행렬과 검출된 개체의 2차원 영상 좌표를 필요로 한다. 따라서 실험 영상은 최종 개체 매칭 이전에 독립적인 두 가지 수행 과정을 거치게 된다. 하나의 과정은 기초행렬을 계산하기 위한 과정으로, 촬영된 두 개의 영상에서 동일 프레임을 이용하여 특징점을 추출하는 것으로 시작한다. 다양한 특징점 추출 알고리즘이 존재하며 각기 다른 장단점을 지니기 때문에 가장 적합한 알고리즘을 확인하여 적용한다. 추출된 특징점들에 등극선 기하학을 적용함으로써 두 카메라와 특징점들 간의 기하학적 상관관계를 나타내는 기초행렬을 계산하면 해당 과정이 종료된다. 다른 하나의 과정은 건설 인력 검출이다. 영상은 2차원 픽셀 데이터로 구성되어 있기 때문에 건설 인력을 검출하기 위한 특별한 알고리즘이 필요하다. 픽셀 그래디언트를 이용하여 형상의 특징 데이터를 생성하고, 기계학습 기법을 적용하여 그것들을 분류함으로써 건설 인력을 검출하면 해당 과정이 종료된다. 두 과정이 모두 종료되면 최종적으로 기초행렬과 검출된 건설 인력 개체의 2차원 영상 좌표를 이용하여 등극선을 계산하면 개체 매칭을 실행할 수 있다.

촬영된 양안 영상에서 각각 독립적으로 특징점을 추출하여 서로의 특징점을 비교하고, 등극선 기하학을 적용하여 추출된 특징점들로부터 두 카메라의 기하학적 상관관계를 표현

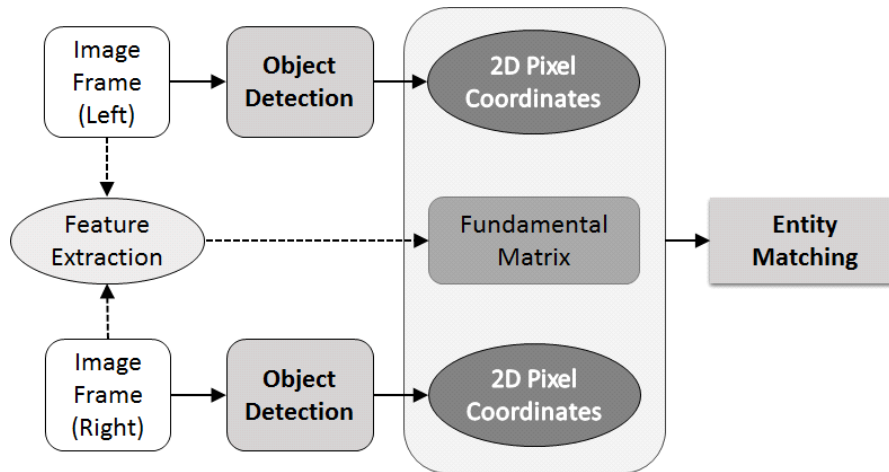


그림 2-27. Epipolar geometry 기반의 Entity matching 프레임워크

한다. 이러한 과정을 통해 두 영상에 등장하는 공통된 부분을 찾을 수 있다. 그렇기 때문에 영상을 매칭하기 이전에, 각 영상에서 특징점을 추출해야 한다.

특징점을 추출하는 것이 중요한 이유는 그림 2-28을 통해 알 수 있다. 그림 2-28의 두 사진에서 육안을 이용해 공통된 부분을 찾다보면 가장 먼저 검은색 명함 상자나 손목의 시계가 쉽게 눈에 띄는 것이다. 하지만 인간이 아닌 프로세서가 그것을 인지하기 위해서는 영상에서 명확하게 확인할 수 있는 특징점들을 필요로 한다. 명확한 특징점을 선택하는 방법은 그림 2-28의 A점과 B점을 비교해보면 알 수 있다. 두 사진에서 A점의 위치가 달라졌음에도 쉽게 해당 점의 위치를 확인할 수 있다. 반면에 B점과 같은 배경은 이동이 없음에도 위치를 정확히 찾아내는 것이 매우 어렵다. 이처럼 영상의 특징을 잘 나타내줄 수 있는 부분을 영상의 특징점이라 한다. 위에서 보았듯이, 특징을 가진 부분이라면 어느 부분이든 특징점이 될 수 있다. 영상 처리에서는 특히 색상, 채도의 변화가 큰 모서리나 꼭지점이 특징점으로 주로 사용된다. 프로세서가 두 개의 카메라로 촬영된 영상들을 비교할 때, 영상의 스케일 변화(scale change), 회전 여부에 따라 동일한 개체를 서로 다른 개체로 인지할 수도 있다. 그러므로 스케일 변화나 회전 여부에 관계없이 정확한 특징점을 추출할 수 있는 적절한 알고리즘이 필요하다. 본 연구에서는 컴퓨터 비전 분야에서 일반적으로 우수하다고 알려진 FAST, SURF와 SIFT 알고리즘 중에서 본 연구에 가장 적합한 알고리즘을 확인하여 적용하고자 한다.

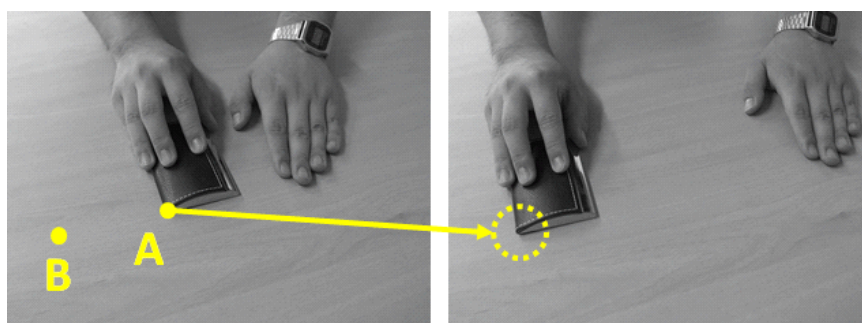


그림 2-28. 특징점 추출의 중요성

FAST(Features from Accelerated Segment Test) 알고리즘은 Drummond 외(2006년)에 의해 제안된 알고리즘으로, 실시간 영상의 중요성이 대두되면서 이를 처리하기 위해 빠른 속도로 특징점을 찾아내도록 고안된 알고리즘이다.

SURF(Speeded-Up Robust Features) 알고리즘은 다중 스케일 공간 정리(Multi-Scale Space Theory)에 기반한 알고리즘이다. 스케일, 조명, 시점 등의 환경 변화에도 불변하는 특징들을 추출해 내는 알고리즘이다. 특징 서술자(Feature descriptor)는 성능과 정확성에서 우수한 헤시안 행렬(Hessian matrix)을 기반으로 검출된다. 성능이 우수하여 일반적으로 널리 쓰이는 SIFT 알고리즘과 비교해서 대등한 성능을 보이는 동시에 속도를 크게 향상시킨 알고리즘이다. 속도를 향상시키기 위해 적분 영상(Integral image)를 추출하여 수행된다.

SIFT(Scale Invariant Feature Transform) 알고리즘은 컴퓨터 비전 분야에서 영상의 특징점을 추출할 때 가장 널리 쓰이는 알고리즘이다. SIFT 알고리즘은 처리 속도가 느리기 때문에 연속적인 영상을 처리하기에는 적합하지 않다. 하지만 개체의 모서리나 꼭지점 등에서 생성되는 특징점을 벡터로 추출하여 영상 회전, 스케일 변화, 유사성을 가진 변형(Affine deformation), 관점 변화(Viewpoint change), 잡음(Noise), 조명 변화(Illumination change)에 대해 강한 특성을 지니며 소개된 알고리즘 중 가장 높은 정확도를 갖는다. SIFT 알고리즘은 (1)스케일 공간 극값 검출(Scale-space extrema detection), (2)특징점 지역화(Keypoint localization), (3)방향성 배치(Orientation assignment), (4)서술자 생성(Descriptor generation)의 4단계 과정을 거쳐 수행된다.

등극선 기하학을 통해 실제 3차원 공간에 위치한 한 점과 그 점이 투영된 두 2차원 영상 평면의 관계를 나타낼 수 있으며 그 상관관계를 표현하는 것이 바로 기초행렬이다. 기초행렬(Fundamental matrix)과 등극선(Epipolar line)의 관계를 그림 2-24를 통해 알 수 있다. 점 x 는 실제 3차원에 존재하는 한 점이며 점 O_L 과 O_R 은 두 카메라의 중심이다. 좌측과 우측에서 같은 점을 촬영하면 각 카메라의 영상에서 점 x 가 투영되는 위치는 서로 다르게 표현된다. 점 x 가 각 영상에 투영된 위치가 x_L 과 x_R 이다. 3차원 상에서 점 x 가 x_1, x_2, x_3 로 이동해도 x_L 의 위치는 변함이 없다. 반면, x_R 의 위치는 점 e_R 과 x_R 을 통과하는 선을 따라 이동하게 되는데 이것을 등극선이라 한다. 다시 말해, 등극선은 실제 점 x 가 투영될 수 있는 후보 위치를 선분으로 표기한 것이다. 그림 2-29에서는 등극선을 l_L 과 l_R 로 표현했다. 기초행렬을 투영 점과 등극선의 관계를 나타내주며 그 관계는 다음과 같다.

$$l_R = F \times L$$

$$l_L = F' \times R$$

기초행렬과 등극선의 특성을 이용하면 2차원 영상에서 검출된 개체와 반대쪽 영상에서 그 개체가 투영될 위치의 후보들을 알 수 있다. 각 영상에서 검출된 개체, 즉 건설 인력의 한쪽 영상에서의 중심 좌표를 알면 기초행렬을 이용하여 반대쪽 영상에서 해당 개체

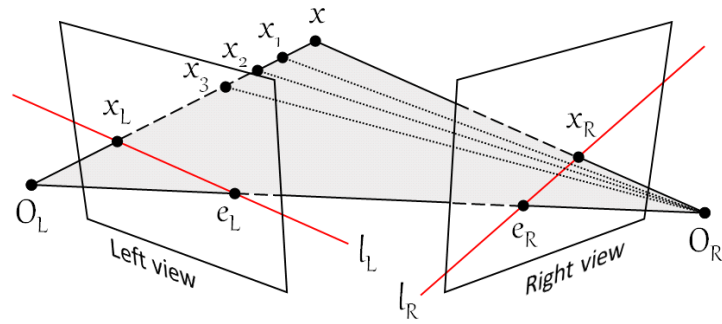


그림 2-29. 스테레오 비전 시스템에서의 등극선

의 위치 후보를 선분으로 표현한 등극선을 얻을 수 있다.

등극선을 이용하면 두 영상에서 독립적으로 검출된 개체를 서로 매칭시킬 수 있다. 그림 2-30의 실제 예를 통해 이것을 쉽게 확인할 수 있다. 좌측 그림은 좌측 영상에서 검출된 개체 중 매칭 목표 개체와 그 중심점을 표기한 것이다. 우측 그림은 같은 지역을 우측에서 촬영한 영상으로, 독립적으로 검출된 각 개체와 그 중심점, 그리고 목표 개체의 등극선을 표기한 것이다. 목표 개체의 중심점은 기초행렬을 통해 우측 영상에서 등극선으로 표현될 수 있다. 정확히 계산된 기초행렬을 토대로 등극선을 계산한다면 등극선은 정확히 한 후보 개체의 중심점을 통과한다. 등극선의 원리에 따라 등극선이 통과한 중심점을 가진 개체가 목표 개체와 일치하는 개체라고 할 수 있다. 하지만 특징점 추출 과정에서 오차와 잡음을 완전히 제거할 수 없기 때문에 실제로 등극선과 중심점은 대부분 정확히 일치하지 않는다.

본 연구에서는 이러한 문제를 제거하기 위해 세 가지 방법을 사용했다. 첫 번째는 등극선과 중심점 사이의 최단 거리에 따른 매칭 방법이다. 잡음으로 인해 등극선이 정확히 매칭 개체의 중심점을 통과하지는 않지만 그 중심점으로부터 등극선까지의 수직거리는 가장 짧을 것이라는 것은 자명하다. 따라서 각 후보 개체들의 중심점과 목표 개체 등극선의 수직거리를 계산하여 그 중 가장 작은 값을 갖는 개체를 매칭 개체로 선정했다. 두 번째 방법은 교차점증이다. 좌측 영상과 우측 영상에서 독립적으로 매칭 과정을 수행하고 두 결



그림 2-30. Epipolar geometry 기반의 Entity matching 결과

과를 종합하여 최종 매칭 결과를 얻어냈다. 세 번째 방법은 반복법을 이용한 기초행렬 계산이다. 본 연구에서는 영상 처리 분야에서 널리 사용되는 반복법 중 하나인 RANSAC(RANdom SAmple Consensus)을 이용했다. RANSAC은 각종 오차, 편차, 노이즈, 이상치가 존재하는 데이터들로부터 최적의 값을 도출하기 위한 절차적 방법을 뜻한다. 무작위로 표본을 추출하여 모델을 생성하고, 모델과 데이터 간의 유사성을 확인하여 유사성이 높은 모델에 가중치를 둔다. 이 과정을 반복하여 가장 지지도가 높은 모델을 최종 모델로 선정하는 것이 RANSAC의 작동 원리다. 어떤 모델을 생성하는데 있어서 가장 적합한 파라미터를 결정해줄 수 있기 때문에 본 연구에서 특징점 매칭 모델을 생성할 때 중요한 파라미터인 기초행렬의 정확한 근사가 가능하다.

(2) 실험 결과

실험 데이터 세트는 실제 건설 현장과 유사한 환경에서 촬영된 영상을 사용했다. 직립 보행 중인 11명의 인원을 대상으로 촬영했으며 동선의 제약 없이 촬영 구역 내에서 자유롭게 보행을 실시했다. 카메라 간의 거리는 2.2m이다. 원본 영상은 MTS 포맷으로 촬영되었기 때문에 소프트웨어가 인식할 수 있도록 H.264 코덱을 이용하여 10fps의 MP4 포맷으로 변환하여 사용했다. 영상의 길이는 총 2분으로 1200개의 프레임을 포함하고 있다. SIFT를 이용한 특징점 추출 연산시간은 76.46초, 기초행렬 계산시간은 0.27초가 소요됐으며 HOG 알고리즘의 총 연산시간은 7570.35초(126.17분), 평균 연산시간은 프레임 당 6.31초가 소요됐다. 이 연구에서는 프로그래밍의 최적화는 이루어지지 않았으며, 추후 멀티스레딩을 통해 멀티코어 CPU(Central Processing Unit)와 GPU(Graphic Processing Unit)를 최대한 활용하는 병렬 프로그래밍으로 연산시간을 크게 단축시킬 수 있을 것으로 사료된다.

다양한 특징점 추출 알고리즘의 효율성을 확인하기 위해, 시각적 방향, 거리, 회전에 따라 차이가 발생한 두 개의 영상을 이용하여 실험을 진행했다. 본 실험에서는 FAST, SURF, SIFT 알고리즘을 적용하여 특징점을 추출했다. 실험에 사용된 영상은 촬영된 영상 중 10개 프레임을 무작위로 선정하여 해상도 변환 없이 PNG 포맷으로 저장된 파일을 사용했다.

양안 영상에 대해 세 개의 알고리즘을 적용하여 10개 영상에서 추출된 특징점의 평균 개수와 매칭된 특징점의 평균 개수, 추출에서 매칭까지의 평균 연산 시간을 표 2-8에 정리했다. 빠른 연산을 수행하도록 고안된 FAST는 연산 시간 면에서 가장 우수했다. 하지만 SURF, SIFT와 비교했을 때, 추출된 특징점의 개수와 매칭된 특징점의 개수가 훨씬 적은 것으로 나타났다. SURF 알고리즘은 FAST 알고리즘보다 연산 시간이 두 배 정도 늘었지만 매칭된 특징점의 개수는 약 9배가량 많다. SIFT는 연산 시간으로 평균 1분 정도를 소비하는 것으로, FAST, SURF에 비해 연산 속도가 매우 느린 것을 확인할 수 있다. 하지만 다른 두 알고리즘과 비교하여 5~7배가량의 특징점을 추출했으며, FAST보다는 20배, SURF보다는 2.5배 많은 매칭 결과를 갖는 것으로 나타났다.



그림 2-31. Entity matching이 정확하게 이뤄진 결과 프레임

본 논문에서 추출된 특징점들은 기초행렬을 계산하기 위해 사용되며 데이터가 많을수록 더 정확한 기초행렬을 계산할 수 있다. 또한, 고정된 카메라를 사용하는 본 논문의 특성상, 특징점 추출과 기초행렬 계산은 최초 1회만 실행되기 때문에 연산 시간은 중요한 고려사항이 아니다. 따라서, 연산 속도가 느리지만 다수의 특징점 데이터를 확보할 수 있는 SIFT 알고리즘이 본 논문에 가장 적합하다고 판단하여 특징점 추출 알고리즘으로 SIFT 알고리즘을 적용했다. 전체 영상 중 첫 프레임에서 SIFT를 이용하여 추출한 특징점들을

표 2-8. FAST, SURF, SIFT 성능 비교

Algorithm	Feature points (Left image)	Feature points (Right image)	Matched points	Computation Time (ms)
FAST	1899.5	2837.7	43.2	275
SURF	2781.8	3070.7	365.4	502
SIFT	14232.3	16799.7	886.9	60,109

이용하여 기초행렬을 계산했다. 특징점은 좌측 영상에서 13906개, 우측 영상에서 16025개를 추출했으며 RANSAC을 통해 잡음과 이상치 데이터를 제거하여 각 영상 당 71개(총 104개)의 특징점을 사용했다.

그림 2-26은 개체 매칭을 수행한 결과들 중 일부이다. 동일 개체로 분류되면 같은 색상의 박스와 레이블로 표시된다. 그림 2-26는 6개의 개체가 모두 정확한 짝을 이루어 매칭된 것을 확인할 수 있다.

표 2-9. Entity matching의 정확도

Number of Processed Frames	Matched Entities			Matching Accuracy
	True	False	Total	
1200	1727	305	2032	85.0 %

표 2-9에 검출 개체 매칭 결과를 정리했다. 총 1200개의 프레임을 대상으로 영상 기반 개체 매칭을 수행한 결과, 총 2032개의 개체가 매칭된 것으로 확인됐다. 이 중, 바르게 매칭된 개체는 1727개, 잘못 매칭된 개체는 305개로 확인되었으며, 매칭 정확도는 약 85%로 계산되었다. 15% 해당하는 잘못된 매칭 결과는 이후 삼각측량에 적용될 경우, 3차원 위치 데이터의 큰 오차로 이어질 수 있다. 따라서, 추후 인력의 형상 및 색상 데이터 등을 활용한 매칭 알고리즘을 추가적으로 도입하여 정확도를 보다 향상시킬 필요가 있으며, 등극선과 개체 중심 좌표의 거리가 일정한 값 이상일 경우 매칭의 결과를 무시하는 방안도 고려해볼 수 있을 것이다. 궁극적으로 안정적인 3차원 위치 추적 시스템을 위해서는 100%에 가까운 정확도를 갖도록 보완될 필요가 있다.

등극선은 목표 개체가 위치할 수 있는 모든 후보 위치를 다른 영상에서 연속적인 직선으로 나타낸다. 그렇기 때문에 하나의 등극선 위에 위치한 개체를 모두 동일 개체로 인식할 수 있다. 오차, 잡음 등에 의해 등극선과 개체의 중심점이 정확히 일치하는 경우는 매우 드물며, 다른 후보 개체가 실제 목표 개체보다 더 가까운 것으로 나타난 경우, 정확한 개체 매칭에 실패할 수 있다. 이 외에도 잡음 개체가 검출되거나 개체가 전혀 검출되지 않아 매칭이 이루어지지 않는 경우도 있다. 따라서 등극선의 일정 범위 내에 여러 개의 개체가 존재할 경우 이들 중 정확한 매칭을 구분할 수 있는 추가적인 알고리즘에 대한 개발이 필요하다.

다. Distance ratio threshold

(1) 목적 및 필요성

Entity matching 과정을 통해 3차원 위치 추적을 위한 개체 매칭 결과를 얻을 수 있지만 보다 안정적인 3차원 위치 추적 시스템을 구현하기 위해 정확도를 향상시킬 수 있는 방법으로 Distance ratio threshold를 적용했다. 표 2-9에서 확인할 수 있듯이 어떠한 threshold도 도입하지 않은 Entity matching의 결과는 85%로 안정적인 3차원 위치 추적

을 위해서는 부족한 정확도를 갖는다.

등극선에 기반한 방법만을 사용하는 것은 정확한 개체 매칭에 한계가 있다. 계산상 발생할 수 있는 각종 오차들로 인해 목표 개체가 반드시 등극선 위에 정확하게 위치할 가능성이 매우 적기 때문이다. Entity matching 과정에서는 해당 등극선과 직선 거리상 가장 가까운 개체를 매칭시키는 방법을 사용하여 이 문제를 해결했으나 이 방법에도 한계는 존재한다. 하나의 선으로 나타나는 등극선의 특성상 해당 선 위에 위치할 수 있는 후보 개체들의 수가 무수히 많기 때문이다. 따라서, 하나의 등극선 주변에 위치하는 여러 후보 개체들 중 가장 높은 신뢰도를 갖는 개체를 매칭시키는 방법으로써 Distance ratio threshold를 적용했다.

(2) 정의 및 적용 방법

Distance ratio는 등극선과의 수직 거리가 첫 번째로 가까운 개체와 두 번째로 가까운 개체의 거리비를 나타낸다. 즉, 하나의 등극선에 매칭될 수 있는 여러 후보 개체들 중 신뢰도가 높은 상위 두 개체를 비교하는 지표이다.

그림 2-32와 식으로 Distance ratio의 정의를 표현했다. 하나의 등극선 주변에 위치한 두 개체 A와 B 중, 어떤 것이 올바른 개체인지 판단하기 위해서는 더 높은 신뢰도를 갖는 개체를 선택할 수밖에 없다. Entity matching 과정에서는 그 신뢰도를 등극선과 해당 개체의 중심점 사이의 수직 거리를 이용했다. 그림 2-27에서는 그 거리를 각각 d_1 , d_2 로 표기했다. Distance ratio threshold가 적용되지 않은 Entity matching 과정에서는 더 작은 수직 거리를 갖는 A 개체가 매칭될 것이다. 하지만 실제로는 B 개체가 옳은 매칭이라고 가정했을 때, 각종 오차들로 인해 d_1 보다 d_2 값이 더 커진 경우라면 해당 매칭은 틀린 결과를 가져오게 된다. 이런 경우를 방지하기 위해 신뢰도의 비를 확인하여 일정 이상의 비를 갖는 경우에만 매칭 결과를 수용한다면 틀린 결과가 수용되는 것을 방지할 수 있을 것이다.

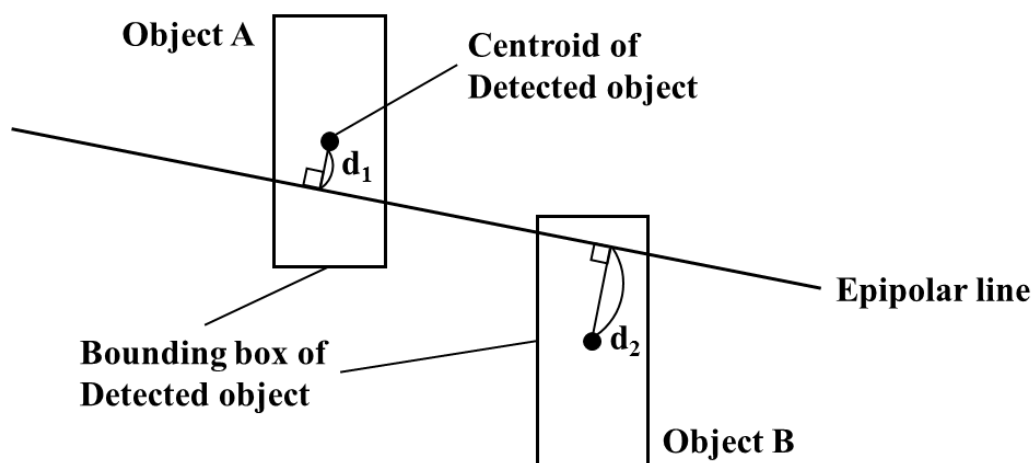


그림 2-32. Distance ratio의 정의 및 그 관계

$$Distance\ ratio = \frac{d_2}{d_1}$$

Distance ratio threshold는 위에 서술한 Distance ratio를 이용한 thresholding 방법으로, 사용자가 지정한 Distance ratio보다 작은 경우, 그 매칭 결과를 기각하는 방식이다. d_1 과 d_2 의 차이가 크지 않을 때, 즉 d_2/d_1 가 1에 가까울수록 두 개체가 갖는 신뢰도는 동점에 가까워진다. 그렇기 때문에 Distance ratio가 크면 클수록 두 개체가 갖는 등극선과의 수직 거리의 차이가 크다는 것을 뜻하며 이것은 A 개체가 갖는 신뢰도가 B 개체에 비해 상대적으로 높다는 것이다. 높은 신뢰도를 갖는 개체와 매칭이 이루어질 경우, 옳은 개체 매칭이 이루어질 가능성이 높으며 또한, 신뢰도가 큰 차이를 갖지 않는 경우, 그 결과를 기각한다면 최종적으로 정확도를 향상시킬 수 있다.

따라서 Distance ratio threshold가 적용된 Entity matching 과정은 다음의 순서를 따라 이루어진다.

- ① 하나의 등극선 주변에 등극선과의 수직 거리가 Distance threshold보다 작은 개체가 한 개인 경우, 등극선과 추적 목표 개체의 중심점 간의 수직 거리가 Distance threshold보다 작은 경우, 개체는 매칭된 것으로 간주한다.
- ② 하나의 등극선 주변에 등극선과의 수직 거리가 Distance threshold보다 작은 개체가 두 개 이상인 경우, 상위 두 개체의 Distance ratio가 너무 작으면 두 개체 모두 매칭에서 제외된다.
- ③ 하나의 등극선에 대해 모든 추적 목표 개체의 중심점을 위의 과정에 따라 확인한다.
- ④ 왼쪽 영상과 오른쪽 영상의 모든 등극선에 대해 위의 과정을 반복하고 서로 같은 결과가 나온 개체는 매칭된 것으로 최종 결정한다.

(3) 실험 결과

실험은 Distance ratio threshold를 1.1부터 4.0까지 0.1씩 높이며 적용한 경우와 전혀 적용하지 않은 경우로 나누어서 진행했다. 결과의 정확도를 확인하기 위해 Recall과 Precision을 계산했다.

표 2-10. Distance ratio threshold를 적용한 결과

Distance ratio thresholding (Ratio)	Recall (%)	Precision (%)
Not Applied	91.21	96.49
Applied (2.5)	80.72	98.10
Applied (4.0)	71.40	98.65

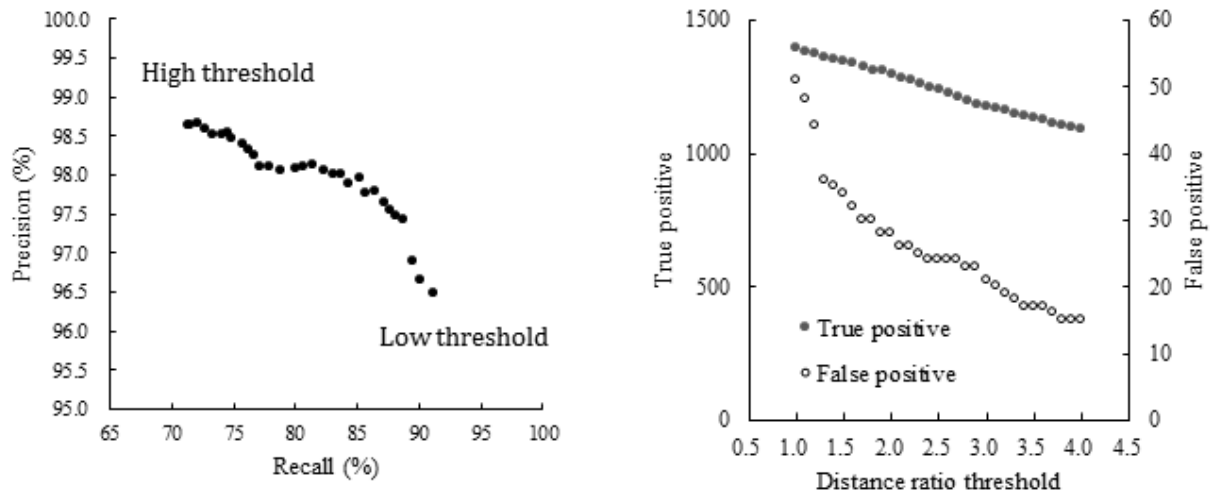


그림 2-33. Distance ratio threshold를 적용한 결과

그림 2-33과 표 2-10은 Distance ratio threshold를 적용하여 특정 영상에서 Entity matching을 과정을 실험한 결과이다. 그림 2-33은 실험 중 설정했던 Distance ratio threshold의 전체 결과를 그래프로 정리한 것이며 표 2-10은 그 중 Distance ratio threshold가 적용되지 않았을 때와 값이 2.5, 4.0으로 설정되었을 때의 결과이다. 전체적으로 Distance ratio threshold의 변경에 따라 Precision과 Recall 사이의 트레이드오프가 있는 것을 확인할 수 있다. 표 2-10을 보면 그 차이를 조금 더 확실히 알 수 있다. Threshold가 적용되지 않은 경우, 높은 Recall과 높은 Precision을 갖지만 3차원 위치 추적에 적합한 Precision으로 보기에겐 어려우며 더 향상될 여지가 남아있다. Distance ratio threshold를 적용했을 때, Recall은 크게 낮아졌지만 향상된 Precision을 확인할 수 있으며 실험 결과 4.0에서 최대의 Precision을 나타내는 것으로 확인되었다. 본 연구에서는 시스템의 검증을 위해 매 프레임 매칭 과정을 수행하지만 제 3차원 위치 추적 시스템은 최초 1회의 매칭만을 필요로 하기 때문에 Recall보다는 Precision에 조금 더 중점을 두고 있다. 따라서, 해당 실험에 사용된 영상에서는 4.0의 Distance ratio를 threshold 값으로 사용했다.

라. Frame grouping과 Motion direction

(1) 목적 및 필요성

3차원 위치 추적 시스템에서 Entity matching 과정은 카메라가 촬영 중인 지역에 건설 인력이 진입했을 때, 최초 1회만 수행되고 이후에는 지속적인 Tracking을 통해 개체를 추적하게 된다. 따라서 Entity matching 과정에서 높은 정확도를 얻을 수 있는 알고리즘을 사용하는 것이 효과적이다. 또한, 초당 25~30프레임으로 이루어진 영상에서 약 30번의 Entity matching 과정이 이루어진다고 해도 불과 1초의 시간만을 필요로 하기 때문에 Entity matching에서 Recall은 큰 고려사항이 아니다. 이러한 이유로, 기존의 Entity matching 방법을 보완할 수 있는 몇 가지 데이터 후처리 방법들을 적용했다. 해당 방법들을 이하에서 각각 Frame grouping과 Motion direction이라 한다.

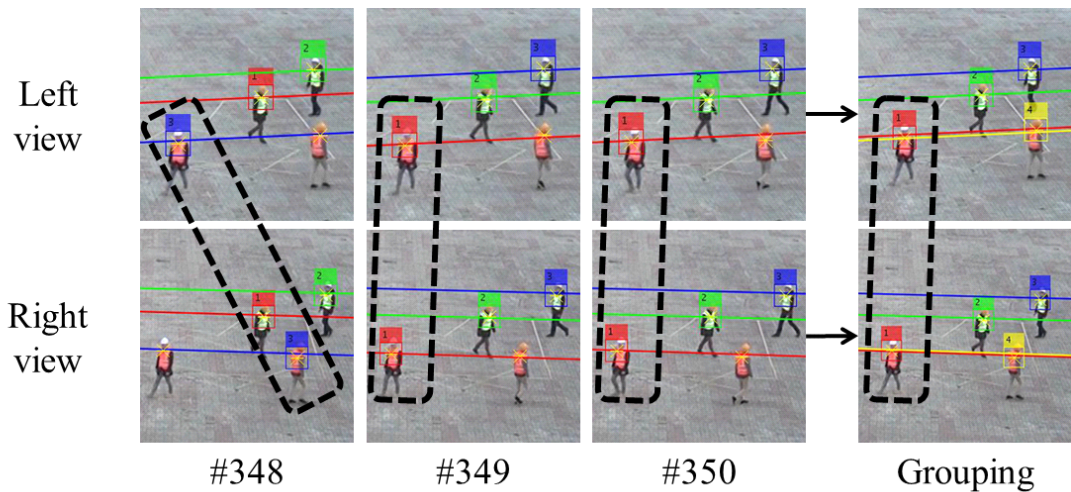


그림 2-34. Frame grouping이 적용된 경우

(2) Frame grouping

Frame grouping은 몇 개의 프레임에 걸친 결과를 종합하여 최종적인 의사 결정을 내리는 방법이다. 사용자가 지정하는 프레임의 개수만큼의 결과를 저장하고 저장된 결과들을 종합하여 다수결에 따라 최종적으로 어떤 개체들이 매칭될지를 결정한다. 이 방법을 적용하면 최종적으로 다수의 결과를 최종 결과로 선택하기 때문에, 비교적 수가 적은 틀린 결과들을 제거할 수 있기 때문에 더 높은 정확도를 기대할 수 있다.

그림 2-34는 Frame grouping이 적용됐을 때의 얻을 수 있는 결과를 나타낸 것이다. Frame grouping이 적용되지 않았을 경우, 3개의 결과 중 #348의 결과는 왼쪽 영상과 오른쪽 영상에서의 동일 인력이 다른 인력과 매칭된 것을 확인할 수 있다. 이처럼 Entity matching에서 틀린 결과를 주게 되면 최종적으로 3차원 위치 추적 과정에서 정확한 결과를 얻을 수 없게 된다. 하지만 Frame grouping을 적용하여 3개 프레임을 결과를 종합하여 최종적으로 결과를 내리는 경우, #348의 틀린 결과는 #349와 #350의 결과에 의해 제거될 수 있다. 이 예시에서는 총 3개의 프레임만을 사용하였지만 사용자의 설정에 따라 원하는 개수만큼의 프레임을 지정할 수 있기 때문에 더 많은 결과를 종합하여 더 높은 정확도를 얻을 수도 있다.

기존의 Entity matching 방법은 한 번의 의사 결정을 통해 개체가 매칭 되기 때문에 최초 결과가 틀린 경우에는 최종적으로 3차원 위치 추적 과정에서 올바른 결과를 줄 수 없다. Frame grouping은 위에 서술한 바와 같이, 다수결에 기반한 후처리 방법이기 때문에 기존의 Entity matching 방법이 갖는 단점을 보완할 수 있다.

Frame grouping은 사용자의 설정에 따라 원하는 개수만큼의 프레임을 사용할 수 있다. 하지만 다수결 방식이기 때문에 짝수로 설정하게 될 경우 동점 상황이 발생할 수 있으므로 홀수로 설정하는 것이 타당하다. 따라서 설정 가능한 최소 프레임 개수는 3개이다. 사용자가 Grouping 하고자하는 프레임의 개수를 설정하면 Entity matching 과정의 결과는

일시적으로 저장된 후 처리된다. 사용자가 5개의 프레임을 설정했다면 추적 목표 개체가 촬영 지역에 진입하여 시스템 상에서 Detection된 이후부터의 Entity matching 결과를 임시로 저장하고 지정된 프레임만큼의 결과가 모이면 이를 토대로 다수결에 따라 자동적으로 최종 결과를 결정하여 매칭이 이루어진다.

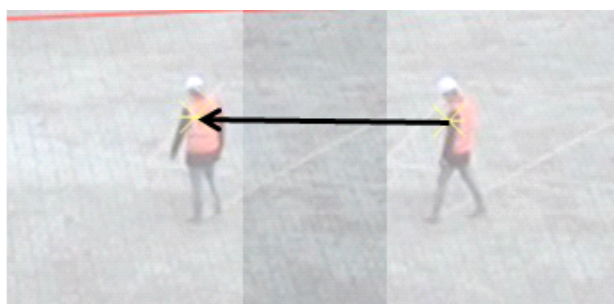
(3) Motion direction

3차원 위치 추적 시스템은 거리 정보를 얻기 위해서 두 대의 카메라를 사용해야만 한다. 그렇기 때문에 화각에 따라 좌우 영상에서 촬영된 개체들의 모습은 조금씩 상이할 수 있다. 촬영하는 대상이 가까울수록 카메라 간의 거리가 멀어지면 각 영상에 투영된 피사체의 모습은 크게 달라진다. 너무 먼 경우에는 피사체의 이동방향이 각 영상에서 다르게 나타날 수도 있다. 하지만 본 연구에서는 비교적 먼 거리에 있는 피사체들을 촬영하고 두 카메라 간의 거리가 그에 비해 크게 멀지 않기 때문에 각 영상에 투영된 피사체의 이동방향이 달라지는 경우는 매우 드물다. 이러한 특성을 이용하여 사용자가 설정한 몇 개의 프레임 사이에 개체가 이동한 방향에 따라 개체의 매칭 여부를 판단하는 방법이 Motion direction이다. 개체들의 이동 방향이 서로 다르다고 가정했을 때, 각 개체의 이동 방향은 개체를 구분 짓는 정보가 될 수 있다. 예를 들어, 두 개의 개체 A, B가 서로 다른 방향으로 이동 중이라면 영상에 투영된 개체들의 이동방향만으로도 개체 매칭을 이룰 수 있다. 하지만 위에 서술한 바와 같이 영상에서 개체들의 이동 방향은 화각에 따라 조금씩 차이가 있을 수 있기 때문에 이동방향만으로 개체 매칭을 이루는 것은 효과적이지 않다. 따라서 기존의 Entity matching에서의 결과를 보완할 수 있는 후처리 과정에서 개체의 이동 방향 정보를 사용하기로 결정했다. 그림 2-35은 Motion direction이 어떻게 적용되는지를 보여주는 예로, #159와 #162 프레임 사이에 개체가 이동했음을 보여준다. 이동한 개체의 중심점 사이의 벡터를 계산하면 해당 개체가 어느 방향으로 이동했는지 알 수 있다. 예에서는 개체가 왼쪽으로 이동한 것을 확인할 수 있다. 이를 왼쪽과 오른쪽 영상에서 각각 수행하여 비교하고, 매칭된 두 개체의 이동방향이 일치하지 않는 경우에는 결과에서 제외시킨다.

사용자는 개체의 이동방향을 계산할 빈도를 설정해줘야 한다. 설정된 빈도에 따라 개체의 이동방향을 계산하고, 그 결과를 이용하게 된다. 너무 낮은 빈도를 설정하게 되면 개체의



#159¢overlapped image



The object moving LEFT

그림 2-35. Motion direction의 적용 예

정확한 이동방향을 계산할 수가 없기 때문에 이를 고려하여 적절한 빈도를 설정하는 것이 효과적이다.

(4) 실험 결과

(가) Frame grouping 실험 결과

실험은 총 1197개 프레임으로 이루어진 영상을 사용했다. Frame grouping의 적용성을 검토하고자 기존의 Distance ratio threshold이 적용된 경우의 결과와 비교했으며 Distance ratio threshold는 가장 높은 Precision을 갖는 4.0을 적용했다. Grouping 개수는 총 3개부터 15개까지 홀수만을 이용했다. 표 2-11에 Frame grouping이 적용된 경우와 적용되지 않은 경우의 결과를 정리했다. 전체 결과 중 13개의 프레임을 이용한 결과가 가장 높은 Precision을 보였으며 이를 적용 전의 결과와 비교했다. Frame grouping이 적용되지 않은 경우 얻을 수 있는 매칭된 개체의 수는 총 1197개의 프레임 중 1111개이지만 Frame grouping이 적용된 경우 전체에서 120개의 결과만을 얻을 수 있었다. 하지만 Precision에서 조금 더 높은 결과를 볼 수 있었으며 99.17%로 100%에 근접한 결과를 보였다. 최초 1회의 매칭에서 정확한 결과를 얻을 수 있다면 최종적으로 3차원 위치 추적 과정에서 조금 더 우수한 결과를 얻을 수 있기 때문에 Frame grouping의 결과가 3차원 위치 추적 시스템에 더욱 적합하다고 판단한다.

표 2-11. Frame grouping의 적용 전과 후의 결과

Frame Grouping (Frequency)	Recall (%)	Precision (%)	Correct matched
Not Applied.	71.40	98.65	1111
Applied (13)	99.17	99.17	120

(나) Motion direction 실험 결과

Motion direction 과정에서도 Frame grouping 과정과 마찬가지로 1197개 프레임으로 이루어진 영상을 사용하여 실험했으며 Distance ratio threshold는 4.0을 적용했다. 실험은 2개 프레임을 이용하는 것부터 10개 프레임까지 1개 프레임씩 늘려가며 수행했다. 실험 결과 2개 프레임 이상부터는 100%의 Precision을 보였으며 프레임의 간격을 늘릴수록 Recall만 낮아지는 경향을 보였다.

표 2-12. Motion direction의 적용 전과 후의 결과

Motion Direction (Frequency)	Recall (%)	Precision (%)	Correct matched
Not Applied.	71.40	98.65	1111
Applied (2)	69.98	100.00	793

최종적으로, Frame grouping과 Motion direction의 결과는 3차원 위치 추적 시스템에 적용하기에 적합한 정확도를 보여준다. 하지만 Motion direction을 적용하는 것이 조금 더 높은 정확도를 얻을 수 있으며 조금 더 빠른 매칭 결과를 얻을 수 있을 것으로 생각된다.

5. 3차원 위치 좌표 계산

가. 카메라 캘리브레이션과 Triangulation

개체 매칭 과정의 결과로 왼쪽과 오른쪽 영상에서 추적하고자 하는 목표 개체의 2차원 영상 좌표를 얻을 수 있으며 해당 영상 좌표들에 삼각측량을 적용하여 개체의 3차원 위치 좌표를 계산할 수 있다. 하지만 실제 영상은 사용된 렌즈, 렌즈와 이미지 센서와의 거리, 렌즈와 이미지 센서가 이루는 각도 등 카메라 내부의 기구적인 부분에 의해서 크게 영향을 받는다. 따라서 2차원 영상에 투영된 좌표들로부터 3차원 공간 좌표를 복원할 때에는 이러한 내부 요인을 제거해야만 정확한 계산이 가능하다. 이러한 내부 요인의 파라미터 값을 내부 파라미터라고 하며, 이것들을 구하는 일련의 과정을 카메라 캘리브레이션이라 한다. 두 대의 카메라를 사용하는 스테레오 카메라 시스템에서는 이런 내부 요인들 외에도 카메라의 자세와 위치를 나타내는 외부 파라미터의 값도 필요하다. 본 연구에서는 3차원 위치 추적을 위해 스테레오 카메라 시스템을 이용하므로 내부 파라미터, 외부 파라미터를 모두 계산해야 한다. 최종적으로 캘리브레이션이 완료되어 두 카메라의 내부, 외부 파라미터들이 계산되면 이를 이용하여 영상의 2차원 좌표들을 3차원 좌표로 변환할 수 있다.

나. 카메라 캘리브레이션

카메라 캘리브레이션은 내부 파라미터와 외부 파라미터를 구하는 방법으로 나뉘어진다. 내부 파라미터는 주점, 초점거리, 비대칭 계수를 포함하는 파라미터로써 카메라의 기계적, 물리적 특성을 나타낸다. 외부 파라미터는 두 카메라 사이의 기하학적 상관관계를 설명하는 파라미터로써, 회전 및 평행이동 변환으로 표현된다.

내부 파라미터는 카메라로 촬영된 이미지들을 이용하여 계산할 수 있다. 일반적으로 정확한 캘리브레이션을 위해 그림 2-36과 같은 동일한 크기를 가진 체크무늬가 그려진 체커보드를 사용한다. 체커보드를 다양한 각도에서 촬영하여 영상을 저장하고, 캘리브레이션 과정에서 각 영상에 나타난 체커보드의 2D 좌표들을 검출하여 카메라의 내부 파라미터를 계산한다. 외에도 카메라의 내부, 외부 파라미터를 계산하는 방법은 다양하지만 본 연구에서는 세 가지 방법에 대해 고려했다.

(1) 내부, 외부 파라미터를 동시에 계산하는 방법

체커보드를 이용하여 고정된 스테레오 카메라를 동시에 캘리브레이션하는 방법이다. 카메라의 내부 파라미터와 외부 파라미터를 동시에 계산할 수 있으며 체커보드의 체크무늬의



그림 2-36. SIFT를 이용한 두 영상에서의 동일 특징점 추출

크기를 알고 있다면 스케일 계수도 바로 계산이 가능하다. 가장 간편하게 캘리브레이션할 수 있는 방법이지만 양쪽 카메라에서 동시에 체커보드가 표시되어야 하기 때문에 카메라와 피사체 사이의 거리가 너무 멀거나 두 카메라 간의 거리가 먼 경우에는 사용하기 어렵다. 또한, 내부, 외부 파라미터를 동시에 계산하기 때문에 캘리브레이션이 완료된 카메라를 이동하게 되면 전체 캘리브레이션 과정을 다시 진행해야 한다.

(2) 내부 파라미터와 외부 파라미터를 각자 계산하는 방법

이 방법에서는 카메라의 내부 파라미터를 계산할 때만 체커보드가 사용된다. 위의 방법과 달리, 두 카메라에서 캘리브레이션 과정을 따로 진행한다. 체커보드를 이용해 각 카메라의 내부 파라미터만 먼저 계산한 뒤에 카메라를 설치할 수 있기 때문에 카메라의 이동이 비교적 자유롭다. 외부 파라미터 계산은 설치된 카메라로부터 촬영된 영상을 이용하여 진행된다. 그림 2-36은 외부 파라미터를 계산하기 위해 SIFT 알고리즘을 이용하여 동일한 특징점들을 추출한 결과이다. 그림과 같이 동일 지역을 촬영한 두 영상에서 특징점들을 추출하여 각 카메라에서의 영상 좌표의 기하학적 관계를 나타내는 Fundamental matrix를 계산한다. 계산된 Fundamental matrix와 각 카메라의 내부 파라미터, 그리고 추출된 특징점들의 2차원 영상 좌표를 이용하면 외부 파라미터를 계산할 수 있다. 카메라의 내부 파라미터 정보를 저장하고 있으면 카메라가 이동한 후에도 비교적 쉽게 외부 파라미터를 계산할 수 있으며, 체커보드가 양쪽 카메라에 동시에 촬영될 필요가 없기 때문에 내부 파라미터를 계산하는 것도 비교적 간단하다.

(3) 컨트롤 포인트를 이용하는 방법

스테레오 카메라로 촬영된 지역에서 특정 점들의 실제 3차원 좌표를 알고 있다면 체커보드를 사용하지 않아도 캘리브레이션이 가능하다. 그런 특정 점들을 컨트롤 포인트라고 하며, 동일한 지역을 스테레오 카메라로 촬영하면서 촬영된 영역 내에 5개 이상의 컨트롤 포인트가 존재해야 한다. 해당 컨트롤 포인트들의 3차원 좌표와 해당 점들이 각 카메라의 영상에 투영된 2차원 좌표 간의 상관관계를 통해 최종적으로 내부 파라미터와 외부 파라미터를 동시에 계산할 수 있다. 이 방식에서도 (2)의 방식과 마찬가지로 특징점 추출 및

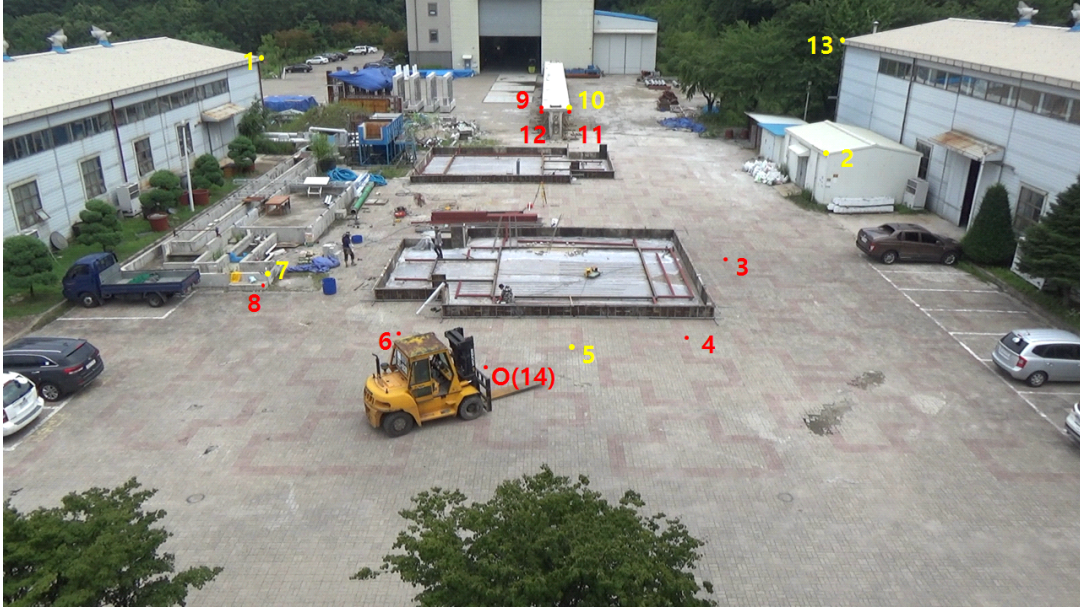


그림 2-37. 예비 테스트 현장 사진 및 Total Station 측정점

매칭을 사용한 Fundamental Matrix의 계산이 필요하다.

본 연구에서는 넓은 영역을 촬영하고 개체의 지속적인 트래킹을 위해 피사체와 카메라 간의 거리가 비교적 멀다. 그에 따라, 충분한 깊이 정보를 알기 위해서 카메라 간의 거리 또한 멀게 설정됐다. 체커보드를 이용하여 내부, 외부 파라미터를 동시에 계산하는 방법은 한 번의 캘리브레이션으로 전체 파라미터를 계산할 수 있다는 장점이 있지만 본 연구의 특성에는 적합하지 않다고 판단했다. 따라서 본 연구에서는 위의 (나)와 (다) 방법을 고려하여 카메라 캘리브레이션을 진행했다.

다. Triangulation

2차원 영상 좌표로부터 3차원 좌표를 계산하는 과정을 Triangulation이라 한다. 본 연구에서는 건설 인력을 추적하는 것이 목적이므로 건설 인력이 각 카메라의 영상에 투영된 2차원 영상 좌표를 사용하여 계산한다. Entity matching 과정에서 얻어낸 목표 추적 개체의 2차원 영상 좌표와 캘리브레이션 과정에서 계산한 외부 파라미터를 이용하면 개체의 3차원 위치 좌표를 계산할 수 있다. 따라서 촬영된 영상의 매 프레임마다 Triangulation을 진행하면 목표 개체의 3차원 위치 좌표를 시간에 따라 추적할 수 있다.

다. Camera Calibration 및 Triangulation 예비 테스트

그림 2-37은 예비테스트를 위해 촬영한 현장 사진과 Total Station으로 3차원 좌표를 측량한 14개 지점을 보이고 있다. 이 중 6 지점을 컨트롤 포인트로 사용하여 5-나-(3)의 방식으로 캘리브레이션하고 triangulation한 결과를 그림 2-38에 보이고 있다. 그림 2-37의 노란색으로 표시된 위치는 컨트롤 포인트로 사용된 점들이고, 빨간점으로 표시된 위치는

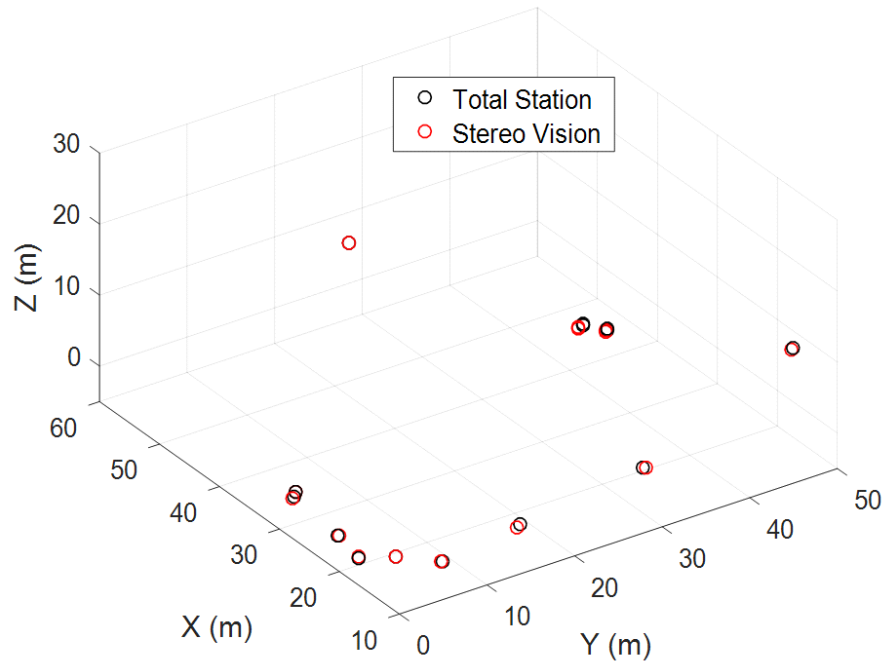


그림 2-38. camera calibration 및 triangulation 예비 테스트 결과

정확도 비교에 사용된 점들이다. 그림 2-39는 Fundamental Matrix 계산에 사용된 특징점 매칭 결과를 보이고 있다.

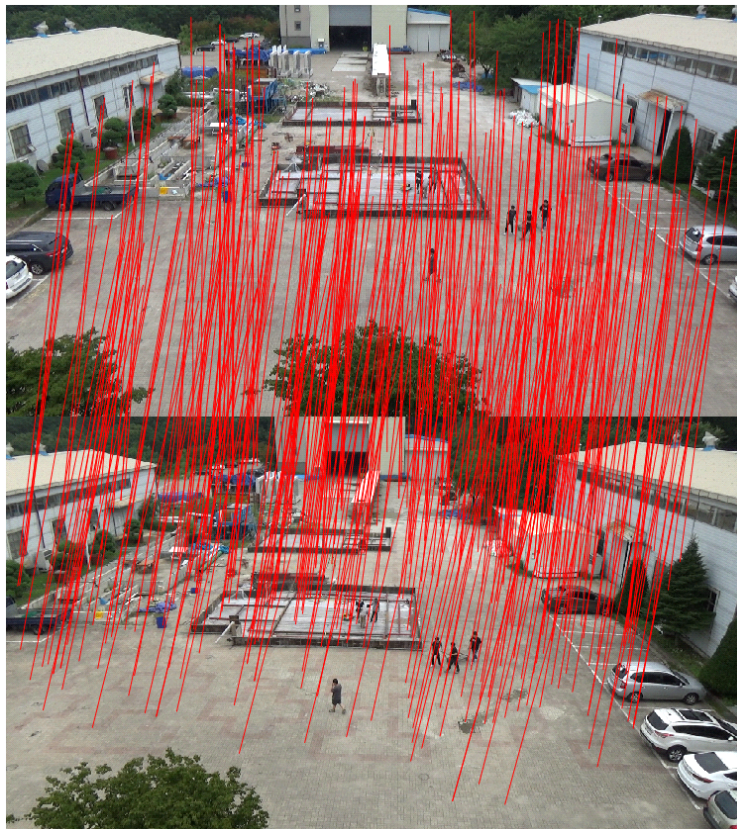


그림 2-39. 특징점 추출 및 매칭 결과

제3장 최종 연구성과 및 적용실적

제1절 국제저널 학술지 게재 논문

1. Yang, J., Park, M.-W., Vela, P.V., and Golparvar-Fard, M. (2015) “Construction performance monitoring via still images, time-lapse photos, and video streams: Now, tomorrow, and the future”, *Advanced Engineering Informatics*, 29, 211-224.
 - 사진 및 영상을 활용한 건설 현장 모니터링 시스템에 관한 연구들을 정리하여 review하는 논문이며, 이 연구 과제에서 수행된 내용이 건설 개체 위치추적에 관한 부분의 주요 내용으로 포함되어 있음.
2. Park, M.-W., Elsafty, N., and Zhu, Z. (2015) “Hardhat-Wearing Detection for Enhancing On-Site Safety of Construction Workers”, *Journal of Construction Engineering and Management*, 141(9): 04015024.
 - 건설인력 탐지와 안전모 탐지 기법을 조합하여 건설현장에서 안전모를 착용하지 않은 인력을 검출하는 방법론에 관한 논문이며, 이 보고서의 [제2장-제2절-1.나]와 [제2장-제2절-2]의 내용이 포함되어 있음.
3. Park, M.-W. and Brilakis, I. (2016) “Continuous Localization of Construction Workers via Integration of Detection and Tracking”, *Automation in Construction* (in 2nd-round review)
 - 건설인력 탐지 기법과 2차원 위치추적 기법을 연계하여 안정적으로 다수의 건설인력을 추적할 수 있는 알고리즘 개발에 관한 논문이며, 이 보고서의 [제2장-제2절-3]의 내용이 포함되어 있음. 현재 2nd-round 리뷰 중에 있는 논문으로 연구성과로 인정받을 수 없지만, 연구책임자 입장에서는 이 연구과제의 성과이므로 여기에 보고함.

제2절 국제 학술회의 발표

1. Zhu, Z., Park, M.-W., and Elsafty, N. (2015) “Automated Monitoring of Hardhats Wearing for Onsite Safety Enhancement”, *Proceedings of the 5th International Construction Specialty Conference (ICSC 2015)*, June 7-10, Vancouver, Canada.
 - [제3장-제1절-2]의 논문이 축약된 내용을 담고 있고 예비실험 결과만을 포함하고 있음. 학술회의에서의 발표를 통해 다양한 의견과 질의를 받아볼 수 있었고, 그 내용 중 일부가 [제3장-제1절-2]의 논문에 반영되었음.
2. Park, M.-W., and Brilakis, I. (2015) “Improved Localization of Construction Workers in Video Frames by Integrating Detection Tracking”, *Proceedings of the 5th International Construction Specialty Conference (ICSC 2015)*, June 7-10, Vancouver, Canada.

- [제3장-제1절-3]의 논문이 축약된 내용을 담고 있음. 학술회의 발표 시 청중으로부터 이 기술의 활용성에 대한 매우 긍정적인 의견들을 들을 수 있었음.
- 3. Park, M.-W., and Zhu, Z. (2015) "Vision-Based Identification of Personal Protective Equipment Wearing", Proceedings of the 6th International Conference on Construction Engineering and Project Management (ICCEPM 2015), October 11-14, Busan, Korea.
- 안전모 또는 안전조끼를 착용하지 않은 건설인력을 자동 검출하는 방법에 관한 논문으로 [제2장-제2절-2]의 마지막 부분의 내용이 포함되어 있음.
- 4. Lee, Y., Park, M.-W., and Brilakis, I. (2016) "Entity Matching across Stereo Cameras for Tracking Construction Workers", Proceedings of the 33rd International Symposium on Automation and Robotics in Construction (ISARC 2016), July 18-21, Auburn, AL.
- 스테레오 비전 시스템에서 두 대의 카메라에서 탐지된 건설인력 개체들을 카메라 간 매칭하는 방법에 관한 논문으로 [제2장-제2절-4]의 내용이 포함되어 있음.

제3절 Software

- 스테레오 영상 데이터로부터 건설인력의 위치를 추적하는 알고리즘들을 .NET Framework 을 기반으로 한 C# 프로그래밍과 Matlab 프로그래밍으로 구현하였으나, 아직 하나의 플랫폼으로 통합되지는 않았으며, 소프트웨어 등록은 이뤄지지 않았음.
- 1. 건설인력 탐지와 2차원 위치추적 알고리즘의 연계, 안전모 미착용자 검출 등의 방법들은 C# 프로그래밍으로 구현하였으며, 단일 카메라 영상을 입력받아 프레임별로 검출 결과들을 디스플레이하는 소프트웨어임. 알고리즘에 내재된 다양한 파라미터들을 소프트웨어 상에서 설정할 수 있으며, 최종 결과물은 텍스트 파일과 프레임별 이미지 파일들로 출력됨.
- 2. 스테레오 비전 기반의 entity matching과 triangulation 부분은 Matlab으로 프로그래밍하여 구현하였으며, Matlab 플랫폼 내에서 동작할 수 있는 스크립트로만 작성되어 있는 상태임. 두 개의 카메라 영상을 입력 받아 매칭 결과를 보여주고 3차원 위치 좌표를 계산함.

제4장 연구목표 달성 및 효과

제1절 연구개발 최종목표 달성도

이 연구에서는 두 대의 카메라 영상을 이용하여 건설인력의 3차원 위치좌표를 계산하는 영상기반 위치추적 시스템을 개발하고자 하였으며, 그림 1-5의 전체 프레임웍 중 굵은 선 테두리로 강조된 부분에 초점을 두고자 하였음.

1. Object detection과 2D tracking 알고리즘의 연계 방안

- 가. 건설인력 탐지 부분에서는 서 있는 자세와 쭈그린 자세의 탐지의 2가지 자세를 탐지할 수 있도록 하였고, 모두 precision 95% 이상의 결과를 보여 목표수준을 달성하였음.
- 나. object detection과 2D tracking의 연계 방안에 대한 연구를 진행하였고, 다양한 실험을 통해 안정적으로 위치추적이 가능한 알고리즘을 개발/구현하였음. 성능평가에서 목표수준을 상회하는 precision 98.9%, recall 98.0%의 결과를 얻었음.
- 다. 상기 내용은 1차년도에 수행 완료하였으며, 주요 연구성과로서 국제 학술회의 발표가 1회 이뤄졌고, 국제 학술지 게재를 위해 논문이 제출되어 현재 리뷰 마지막 단계에 있음. 해당 내용은 단일 소프트웨어로 구현되었음(소프트웨어 등록 신청 전 상태임).

2. Entity Matching

- 가. 1차년도에 epipolar geometry 기반으로만 매칭을 하여 85%의 precision을 얻었고, 여기에 추가 알고리즘들을 도입하여 최종적으로 목표수준을 상회하는 100%의 precision을 얻었음.
- 나. 2차년도에 수행 완료하였으며, 주요 연구성과로서 국제 학술회의 발표가 1회 이뤄졌고, 해당 내용은 Matlab 프로그래밍을 통해 구현되었음.

3. 안전모 및 안전조끼 미착용자 검출

- 가. 건설인력 탐지 알고리즘을 응용/확장하여 건설현장 안전관리 목적으로 안전모 및 안전조끼 미착용자를 자동 검출하는 방법에 대한 연구를 수행하였음.
- 나. 안전모 미착용자 검출은 94.3% precision의 높은 성과를 보였고, 테스트를 확장하여 알고리즘 보완 및 파라미터 최적화를 통해 정확도를 향상시키면 현장 적용 수준에 이를 수 있을 것으로 사료됨.
- 다. 안전모 또는 안전조끼 미착용자 검출은 아직 precision이 86.8%에 그쳐, 지속적인 연구를 통해 정확도 향상을 이뤄야 함.
- 라. 1차년도에 수행 완료하였으며, 주요 연구성과로서 국제 학술회의 발표가 2회 이뤄졌고, 국제 학술지 논문 1편이 간행되었음. C# 프로그래밍 소프트웨어로 구현되었음.

4. Triangulation을 통한 3차원 위치추적

- 가. 현재 연구 진행 중인 부분이며, 3차원 위치추적 데이터의 정확도를 위한 실험이 남아 있는 상태임.

제2절 연구개발 성과의 기술적 효과분석

1. 본 연구의 가치는 저비용으로 건설현장 개체들의 위치추적을 할 수 있는 시스템을 창출한다는 데에 있다. 이 연구 결과가 곧바로 현장 적용이 가능한 시스템 구축으로 이어질 수 있는 수준에 도달하기는 힘든 단계이지만, 현장 적용에 한발 더 가까이 다가가는 단계는 마련할 것이라 생각하며, 일부 간단한 시나리오와 환경에 대하여, 현장 적용도 시험해 볼 수 있을 것이다.

가. 비교적 높은 곳에 카메라를 설치하여 촬영한다면, occlusion 영향을 줄여 적용 가능성을 높일 수 있다. 한정된 특수 시나리오에서 순조롭게 추적 데이터를 얻을 수 있다면, 효율적 건설관리가 부분적으로나마 도모할 수 있을 것이다.

나. 본 연구가 지속적으로 이루어진다면, 생산성 관리, 현장안전 관리 등에 기여할 수 있을 것이다. 이는 결과적으로 현장 관리자가 효율적으로 현장을 관리할 수 있는 기반을 마련해 주리라 기대한다.

(1) 생산성 관리: 현장 구역별/작업별로 동원된 인력의 수, 동선, 활동 범위에 대한 데이터를 위치추적 데이터로부터 얻을 수 있으며, 이러한 정보는 작업별 생산성 관리에 중요한 정보로 활용될 수 있다.

(2) 현장안전 관리:

(가) 건설 현장에 현장안전을 위하여, 일반 건설 인력의 출입이 제한되는 구역이 지정되어 있으므로, 이를 위반하여 발생할 수 있는 안전사고를 위치 추적 데이터를 활용하여 미연에 방지하도록 한다.

(나) 건설 중장비들의 위치 데이터를 GPS를 통하여 수집하는 경우, 건설인력과의 상대적 위치 관계를 통하여 인력과 중장비 간의 충돌 사고를 미연에 방지하도록 한다.

(다) 현장에서의 안전모, 안전조끼 착용여부를 탐지하여 미착용자에 대한 위치정보를 제공하는 간단한 프로그램을 탑재하여 안전관리에 활용할 수 있다.

다. 위치추적 결과의 시각화를 위한 BIM과의 연계시스템도 구현 가능할 것이라 판단된다.

2. 최근 건설의 효율화를 위하여 다양한 센싱 기술들의 현장 도입이 연구되고 있다. 위치추적용으로는 GPS, RFID가 대표적인 기술들이다. 각각이 서로 다른 장단점들을 지니고 있고, 그에 따라 적합한 적용 분야가 서로 다르기 마련이다. 여기에 이 연구에서 제시하고 있는 영상 기반 기술을 더하여, 각 기술들에 대한 종합적 분석을 통하여 세 가지 기술의 서로 다른 활용 방안을 정리하고, 더 나아가 세 기술이 통합된 상호보완 시스템에 대한 구상도 가능하리라 기대해 본다.

3. 영상 기반의 기술들은 저렴한 카메라 장비를 사용한다는 장점이 있으며, 신호처리 알고리즘의 개발로 인간의 시각을 대신하여 작업을 자동화할 수 있다는 데에 효용성이 있다. 건설분야에서도 구조물 진단, 구조물 3D 모델링 등 다각도로 이 기술의 효용성에 대한 연구들이 진행되고 있다. 본 연구의 결과도 이의 한 분야로서 영상의 활용도를 높이는 데에 일

조하고, 학계와 업계의 관심과 지원을 유도하여 보다 활발한 국내 연구활동을 도모할 수 있기를 기대한다.

4. 이 연구는 카메라를 이용한 연구로서 상당 부분의 결과물이 영상으로 표현된다. 이러한 결과를 시각적 교육자료로 활용하여, 토목환경공학과 학생들로부터 건설관리 자동화 분야에 대한 관심을 이끌어 내어 보다 높은 참여도를 도모할 수 있을 것으로 기대된다.

제3절 연구개발 성과의 경제적 효과분석

1. 이 연구에서는 건설 현장 내에서의 건설인력들의 이동 이력을 자동 추적하는 시스템을 개발한다. 실제 현장에서 사용된 바가 있고, 꾸준한 연구가 진행되어온 GPS 및 RFID 기반 추적시스템에 비하여 비용과 시간 면에서 모두 효율적인 시스템을 구축하는 데에 목적을 둔다. 체계적 현장 모니터링에 대한 지속적인 연구는 건설현장 작업의 질과 효율성을 상당한 수준 향상시킬 수 있을 것이라 기대한다. 궁극적으로 건설업계의 경쟁력을 높이고, 사회기반시설의 생애주기 비용을 절감할 수 있는 효과로 이어질 수 있기를 기대한다.
2. 여러 핵심 기술을 습득한 새로운 전문 인력의 양성을 기대해 볼 수 있다. 건설공학 측면에서 국내외에서 그 중요성이 증대되고 있는 건설자동화를 통해 IT 기술과 건설관리 지식을 통합적으로 보유하고, 소프트웨어 제작 및 현장 테스트 과정의 경험까지 갖추고 있는 우수한 인력의 양성이 이루어질 것으로 기대한다. 공학자로서 핵심적인 기초학문과 응용학문을 동시에 경험하고 보유하게 될 것이다.

제5장 연구성과의 활용 및 추가연구 필요성

제1절 연구성과 향후 활용방안

1. 이 연구의 궁극적 목표는 건설현장 인력의 움직임을 실시간으로 파악하여, 현장 시공과정을 모니터링하는 데에 활용할 수 있도록 건설인력용 Tracking 시스템을 개발하는 것이다. 기존의 GPS, RFID 등의 무선주파수 기반 방식을 탈피하여 영상기반 시스템을 제안하고 있다. 무선주파수 기반의 시스템들과의 정확한 비교/분석을 통하여, 각 시스템 별 장단점을 파악하고, 현장조건, 추적개체 종류 등의 조건에 따른 최적의 시스템을 선정한다면, 실제 현장에서 활용 가능한 지침이 될 것으로 생각한다. 무선주파수 기반 시스템들과의 조합으로 서로의 장점을 부각시키면서 단점을 보완하는 시스템 구성을 기대해 볼 수 있다.
2. 최근 빅데이터 분석 연구가 활발해지면서, 그 일환으로 건설현장 개체의 이동경로 및 움직임 데이터에 대한 분석 연구가 이뤄지고 있다. 건설인력 또는 장비들의 시간이력 위치데이터에 관한 패턴을 분석하여 작업 종류 및 능률, 생산성 등을 판단하는 것이다. 이러한 연구를 위해서는 방대한 양의 위치추적 데이터를 수집해야 하므로, 이 연구에서 개발된 건설인력 위치추적 시스템을 연계할 수 있을 것으로 사료된다. 실제 연구 책임자와 지속적 교류를 맺고 있는 영국 캠브리지 대학의 construction IT 연구실에서 유사한 연구들이 진행되고 있어 연구 협력을 모색하고 있다.
3. 최근 웨어러블 디바이스가 보급되면서 건설인력에 활용하는 방안이 검토되고 있다. IMU(Inertial Measurement Unit)와 같은 센서를 이용하여 데이터 패턴 분석을 통해 건설인력의 작업 동작에 대한 정보를 추출하는 것이다. 이러한 정보에 위치정보가 연계되면 보다 정확한 작업 분석이 가능할 것으로 사료된다.

제2절 추가연구 필요성 등

1. 궁극적으로는 건설인력 위치추적 시스템을 현장적용이 가능한 단계로 구현하는 것이 최종의 궁극적 목표이지만, 이 연구에서는 2년의 기간 동안 그 근간이 될 소프트웨어를 개발하는 것을 목표로 하였다. 향후 지속적인 후속연구가 이어진다면, 일정 수준의 안정성과 정확도를 확보할 수 있도록 각 구성단계 별 알고리즘들을 보완하고 이루고, 프로그램을 최적화하여 실시간 프로세스가 가능한 시스템 개발이 가능할 것이라 기대한다. 더 나아가, 건설인력 이외의 개체들(건설장비)의 추적에도 적용 가능한 시스템으로 확장 또한 기대해 볼 수 있을 것이다.
2. 이 연구에서는 2대의 카메라만을 사용하였으나, 3대 이상의 카메라를 사용할 경우 보다 높은 정확도를 기대해 볼 수 있다. 따라서 다양한 형태의 카메라 구성 및 카메라의 최적 위치 선정에 대한 연구 역시 추가적으로 필요한 연구라 할 수 있다.

참고문헌

- Alahi, A., Raphael, O., and Vandergheynst, P. (2012) "FREAK: Fast Retina Keypoint." 2012 IEEE Conference on Computer Vision and Pattern Recognition, 510-517.
- Brilakis, I., Park, M.-W. and Jog, G. (2011a) "Automated vision tracking of project related entities." Journal of Advanced Engineering Informatics, 25(4), 713-724.
- Brilakis, I., Fathi, H., and Rashidi, A. (2011b) "Progressive 3D reconstruction of infrastructure with videogrammetry." Automation in Construction, 20(7), 884-895.
- Chae, S., and Yoshida, T. (2010) "Application of RFID technology to prevention of collision accident with heavy equipment." Automation in Construction, 19(3), 368-374.
- Chi, S. and Caldas, C.H. (2011) "Automated Object Identification Using Optical Video Cameras on Construction Sites." Computer-Aided Civil and Infrastructure Engineering, 26(5), 368-380.
- Caldas, C.H., Torrent, D.G., and Haas, C.T. (2004) "Integration of automated data collection technologies for real-time field materials management." The 21st International Symposium on Automation and Robotics in Construction, Jeju, Korea, Proceedings.
- Costin, A., Pradhananga, N., and Teizer, J. (2012) "Leveraging passive RFID technology for construction resource field mobility and status monitoring in a high-rise renovation project." Automation in Construction, 24, 1-15.
- Dalal, N. & Triggs, B. (2005), Histograms of oriented gradients for human detection, in Proceedings of 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1, 886-893.
- Ergen, E., Akinci, B., and Sacks, R. (2007) "Tracking and locating components in a precast storage yard utilizing radio frequency identification technology and GPS." Automation in Construction, 16(3), 354-367.
- Golparvar-Fard, M., Peña-Mora, F., Arboleda, C. A., and Lee, S. H. (2009) "Visualization of construction progress monitoring with 4D simulation model overlaid on time-lapsed photographs." Journal of Computing in Civil Engineering, 23, 391-404.
- Gong, J. and Caldas, C.H. (2010) "Computer Vision-Based Video Interpretation

- Model for Automated Productivity Analysis of Construction Operations.” *Journal of Computing in Civil Engineering*, 24(3), 252–263.
- Ko, C.H. (2010) “RFID 3D location sensing algorithms.” *Automation in Construction*, 19(5), 588–595.
- Leutenegger, S., Chli, M., and Siegwart, R.Y. (2011) “BRISK: Binary Robust invariant scalable keypoints.” 2011 IEEE International Conference on Computer Vision, 2548–2555.
- Lu, M., Chen, W., Shen, X.S., Lam, H.C., and Liu, J.Y. (2007) “Positioning and tracking construction vehicles in highly dense urban areas and building construction sites.” *Automation in Construction*, 16(5), 647–656.
- Mcfarlane, N.J.B. & Schofield, C.P. (1995), Segmentation and Tracking of Piglets in Images, *Machine Vision and Applications*, 8(3), 187–193.
- Oloufa, A.A., Ikeda, M., and Oda, H. (2003) “Situational awareness of construction equipment using GPS, wireless and web technologies.” *Automation in Construction*, 12(6), 737–748.
- Park, M.-W., Makhmalbaf, A., and Brilakis, I. (2011) “Comparative study of vision tracking methods for tracking of construction site resources.” *Automation in Construction*, 20(7), 905–915.
- Park, M.-W., and Brilakis, I. (2012a) “Construction Worker Detection in Video Frames for initializing Vision Trackers.” *Automation in Construction*, 28, 15–25.
- Park, M.-W., Koch, C., and Brilakis, I. (2012b) “Three-Dimensional Tracking of Construction Resources using an On-site Camera System.” *Journal of Computing in Civil Engineering*, 26(4), 541–549.
- Rezazadeh Azar, E., Dickinson, S., and McCabe, B. (2013) “Server-customer interaction tracker; a computer vision-based system to estimate dirt loading cycles.” *Journal of Construction Engineering and Management*, 139(7), 785–794.
- Rezazadeh Azar, E., Feng, C., and Kamat, V.R. (2015a) “Feasibility of in-plane articulation monitoring of excavator arm using planar marker tracking” *Journal of Information Technology in Construction(ITcon)*, 20, 213–229.
- Rezazadeh Azar, E. (2015b) “Construction Equipment Identification Using Marker-Based Recognition and an Active Zoom Camera” *Journal of Computing in Civil Engineering*, 30(3), 04015033.

- Ross, D., Lim, J., Lin, R.-S., & Yang, M.-H. (2008). Incremental learning for robust visual tracking, *International Journal of Computer Vision*, 77(1), 125-141.
- Swain, M.J. & Ballard, D.H. (1991), Color Indexing, *International Journal of Computer Vision*, 7(1), 11-32.
- Wang, L.C. (2008) "Enhancing construction quality inspection and management using RFID technology." *Automation in Construction*, 17(4), 467-479.
- Woo, S., Jeong, S., Mok, E., Xia, L., Choi, C., Pyeon, M., and Heo, J. (2011) "Application of Wi-Fi-based indoor positioning system for labor tracking at construction sites: A case study in Guangzhou MTR" *Automation in Construction*, 20(1), 3-13.
- Yang, J., Arif, O., Vela, P.A., Teizer, J., and Shi, Z.K., (2010) "Tracking multiple workers on construction sites using video cameras." *Advanced Engineering Informatics*, 24(4), 428-434.

중요 연구 변경사항

1. 안전모 및 안전조끼 미착용자 검출에 대한 연구가 원 제안서에 포함되지 않았던 내용이나, 건설인력 탐지에 관한 연구와 연관성이 높고, 이 연구 본연의 목적인 3차원 위치추적에 비해 현장 활용성이 높은 것으로 판단되어 병행하기로 결정하였음.
2. 반면에, 3차원 위치추적 프레임웍의 단계별 알고리즘들을 하나의 플랫폼으로 통합하는 부분을 배제하였음. 이 부분은 온전히 프로그래밍 작업만을 요하는 부분으로 연구의 성격이 비교적 덜하며 많은 시간이 소요될 것으로 예상되어 비교적 연구기간이 짧은(2년) 이 연구에서 배제하기로 결정함. 이보다는 안전모/안전조끼 미착용자 검출에 관한 연구를 추가 수행하고 통합 플랫폼 제작 없이 바로 3차원 위치 데이터 정확도 검증 부분에 초점을 두는 방향으로 일부 연구 계획을 변경하였음.

주 의

1. 이 보고서는 국토교통부에서 시행한 국토교통기술촉진연구사업의 연구보고서입니다.
2. 이 보고서 내용을 발표하는 때에는 반드시 국토교통부에서 시행한 사업의 연구결과임을 밝혀야 합니다.
3. 국가과학기술 기밀유지에 필요한 내용은 대외적으로 발표 또는 공개하여서는 아니됩니다.