JunYoung Gwak
Junior, Computer Science

Mentored by the
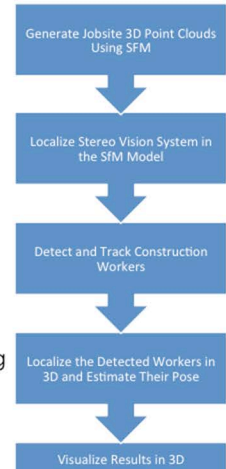Image Spatial Data Analysis group

Project Objective

The goal of this project is to automatically detect and track construction workers and equipment from 2D video cameras and localize their position in 3D purely based on computer vision based technique.

This project will make automated safety monitoring affordable, potentially preventing many accidents on construction sites.

Approach

My work focuses on developing a prototype for automated detection and localization of construction workers in 3D. This involves generating jobsite 3D point clouds using **Structure-from-Motion (SfM)**, localizing stereo streams in the point cloud using **Perspective-n-Point (PnP)**, detecting, tracking, and localizing workers, plus estimating their pose in 3D. The results are visualized in an interactive **web interface** using WebGL.

Generate Jobsite 3D Point Clouds Using SFM

↓

Localize Stereo Vision System in the SfM Model

↓

Detect and Track Construction Workers

↓

Localize the Detected Workers in 3D and Estimate Their Pose

↓

Visualize Results in 3D

Status

1. Produced noise-fee point cloud models from site images;
2. Localized Stereo Vision system in reconstructed point cloud models;
3. Detected construction workers based on **background subtraction**;
4. Localized the detected workers in 3D using **Epipolar geometry**;
5. Visualized the outcome in a web-based environment

← Interactive web interface to visualize the location of the detected workers in 3D point cloud model

↓ Worker detection in 2D using background subtraction

x: -0.229
y: 0.457
z: 4.613

↓ The pipeline of producing 3D image-based point cloud models

Feature Extraction → Feature Matching → Match Tracking → Point Cloud Generation

↓ The pipeline of localizing stereo camera in existing point clouds

Feature Extraction → Feature Matching with 3D Point Cloud → Projection Ray Reconstruction → Pose Estimation

Questions

- How unordered collection of site images can produce accurate 3D point cloud models?
- How movable stereo vision systems could be localized in site coordinate systems?
- How workers can be detected reliably in 2D?
- How detected workers could be localized in 3D point cloud models?
- How to visualize the outcome in real-time interactive 3D environments to enable safety inspection and jobsite decision making processes?