Stereo Vision based Hazardous Area Detection for Construction Worker's Safety

Doyeop Lee[a], Numan Khan[a] and Chansik Park[a]

[a]Department of Architectural Engineering, Chung-Ang University, Seoul, South Korea

E-mail: doyeop@cau.ac.kr, numanpe@gmail.com, cpark@cau.ac.kr

Abstract –

In order to ensure the safety of construction workers during the construction process, supervision of violations of safety regulations and control of hazardous situations not recognized by the workers are continuously required. However, due to the dynamic nature of the construction site, the hazardous situation is always changing and there is a limit to managing all these risk situations in a manual manner by the supervisor. To address these limitations, various computer vision-based researches are being conducted that can automatically identify and manage risk factors. The aim of this study is to propose a system for detecting potential horizontal and vertical risks of moving or static objects around workers by measuring distance between objects using Stereo Vision. Proposed system consists of motion detection, object detection and distance measurement between objects using depth image from stereo camera. It synthesizes images from multiple cameras into one and uses motion-detection and object-detection to identify the movement of objects in a vertical or horizontal position. Then, based on the depth information of the stereo camera, the distance is calculated to identify whether a worker is in the hazardous zone. The proposed system was implemented in the lab environment to verify the proposed functionality. The lab test results indicate that the proposed system can detect vertical and horizontal hazards. It is expected to contribute to the prevention of possible accidents by detecting and controlling the dangers around the workers in real time.

Keywords –

Construction Safety; Depth Measurement; Distance Calculation; Risk Area; Stereo Camera

1 Introduction

Despite various efforts to reduce the number of accidents and fatality in construction sites, construction Safety accidents are occurring continuously. Safety in construction remains a critical issue. Among various types of accidents, this study aims to focus on ways to manage risk situations that are not recognized by workers.

Based on the KOSHA database, the accident cases that happened during 2016-2018 have reported 1345 fatalities in construction. The reported three years of data have been analyzed to find the proximity and collision-related accident cases. As a result, 222 cases are identified related to proximity risks, either horizontally or vertically, that account for 12% of the total fatalities that happened during that interval of time.

For instance, the accident that happened during vertical simultaneous work is a case in which a worker hit by fallen rebar lifted by a crane, hitting by the fallen pipe from the upper floor while working near the window, and hitting by the falling temporary structure while working at the basement. The accident happened during horizontal simultaneous work is a collision happened with a moving car while cleaning up the site, falling into an opening during the site inspection.

These fatal accidents may occur because the worker is not aware of the dangerous situation, but in some cases, the accident is caused by the worker ignoring the dangerous situation. The main purpose of this study is to propose a stereo-vision based system that can determine in real time whether a worker is exposed to a dangerous situation by measuring the distance between the worker and the risk factors. Through this, it is possible not only to control the dangerous situation, but also to identify how often workers are exposed to the danger at the site, and to provide information that can be reflected in future safety management plans.

2 Related Work

2.1 Sensor-based Proximity Risk Detection Systems

The complex nature of the construction site and dynamic movements of construction entities tend to produce many struck-by hazards. To avoid collisions between the objects, the proximity warning system (PWS) is widely used [1]. The PWS operates based on the spatial relationships between the interested entities, and their performance depends on the type of object's tracking technology, for instance: radio frequency (RF) sensing, ultrasonic, radar, global position system (GPS), and magnetic field [2]. To determine the proximity between the objects, the radiofrequency sensing technology emits the electromagnetic signals from a mounted device on objects [3]. This technology uses predefined rules to generate an alert based on signal strengths between the device and objects with the attached tags. Based on the power source in the tags, RF technology can be grouped into three classes:(1) active RF technology and (2) passive RF technology [4]. The active technology needs a power-source for long-range detection. For instance, a type of active RF technology, Ultra-Wideband (UWB) robustly transmits the signals using comparatively low power for tracking the objects, with the errors of centimeter-level [5]. However, passive RF technology requires a power source for the RF device in the shorter-range detection. Passive tags are less expensive, powered by the radio waves from the reader, and have a range of up to 30 feet [4]. Since the RF-based PWS only requires signals strength adjustments and data processing step, thus, it can be easily implemented in the construction site [5]. However, RF technology requires additional work of putting tags on every interested object. Moreover, their signals could be interfered and disturbed by the multipath fading effect [1].

On the contrary, the Global Positioning System (GPS) is employed to recognize the absolute location of objects [3,6]. The GPS sensor identifies the latitudinal and longitudinal coordinates of the objects which required to be monitor. To detect the proximity risk of the entities, the heading direction and speed also need to be considered and extracted [1]. GPS can be used to cover large scale area; thus, many researchers have extensively studied this technology for the practical implementation in the large construction sites [2,7]. However, previous efforts revealed various limitations such as manual input from the supervisor to limit the access to a hazard zone by using risk zone information [8], and accuracies dependent cost ranges of the GPS sensors. Moreover, the additional significant limitation is the strength of the signals, which could be blocked by buildings or other objects.

2.2 Vision-based Proximity Risk Detection Systems

To detect the proximity risks in construction, previous studies have used object tracking techniques for monitoring the trajectories of the localized objects in a series of images [7,9,10]. The primary step for object tracking and action recognition includes the detection of objects. Once the entities of interest are successfully recognized, the location of the bounding boxes in a frame with respect to time could be traced by utilizing the object tracking algorithm [11]. The localization of objects takes place based on their pixel information, changes in brightness, intensity, and other local features. Detection of the sudden slope failure is possible based on the change in pixel intensity [12]. The extracted location information is then used to identify unsafe actions and conditions.

Compared to sensor-based tracking techniques such as RFID tags, UWB, and GPS, Computer vision as a technique has the capability to identify multiple object's information form construction worksites using images only, without human involvement or any additional devices for instance, tags or sensors. Furthermore, the camera can cover large area, thus, multiple objects can be simultaneously captured and easily tracked. Location and detailed information such as types and velocity of the objects can also be acquired. However, the performance of the vision-based systems may vary under poor light conditions, snow, rain, or dusty storms.

3 Proposed Methodology

The purpose of this study is to propose a stereo vision-based risk detection system that can identify how close a worker is to vertical and horizontal hazards and determine whether they are dangerous. Figure 1. illustrates the operating process of the proposed system.

The system consists of motion detection, object detection, and object depth and distance measurement technology using Stereo vision. This system is designed to use the stereo camera's depth calculation in a way that works only under certain conditions, so that it can be used in real time even on low computing resources. First, motion detection and object detection algorithms are applied to the stereo video information to determine the movement and presence of objects in the air and on the ground, and to check the possibility of a dangerous situation. After that, the distance between objects is measured based on the depth information of the stereo camera to determine whether the distance between objects located on the ground is a dangerous situation horizontally and whether an object located in the air poses a danger to objects on the ground.

This system is applied Motion Detection technology that detects motion compared to the current frame and previous frames, YOLO V4 Object Detection algorithm that enables stable real-time detection, and the depth measurement technology of the object using INTEL RealSense Stereo Camera. This stereo camera measures the depth of a specific pixel value through operation using video taken from two or more different angles and additional infrared or RGB color image sensors. When using a general camera, more computation is required at the same resolution, which increases the consumption of computing resources. Thus, the proposed method is designed to minimize depth calculation by utilizing the stereo camera (Intel RealSense) so that it can operate in real time even with low computing resources. For this purpose, motion detection and object detection are designed to operate sequentially in one video information, which identifies the movement and presence of objects in the air and on the ground, and identifies the possibility of dangerous situations. Then, after calculating the distance based on the depth information of the stereo camera, it is determined whether a specific object is in a dangerous situation.

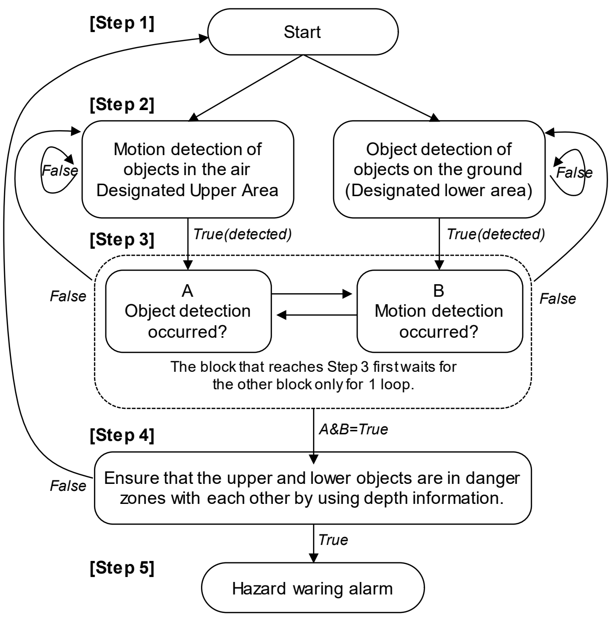

Figure 1. Overview of the proposed process

4 Lab Test and Results

4.1 Lab Test Process

Lab test was conducted on two conditions for determining vertical and horizontal hazards (Figure 2). In the case of horizontal hazards, the distance between a worker working on the ground and a dump truck was measured, and in the case of vertical hazards, the distance between a worker working on the upper area and a worker working on the ground was measured. The risk level was set to be dangerous when the distance between vertically and horizontally located objects is within 2 m, caution when 2 to 4 m, and safe when 6 m or more. For the lab test, a 1/20 scale model was used.

Depth estimation using stereo vision from two images (taken from two cameras separated by a baseline distance) involves three steps: First, establish correspondences between the two images. Then, calculate the relative displacements (called "disparity") between the features in each image. Finally, determine the 3-d depth of the feature relative to the cameras, using knowledge of the camera geometry [13]. The typical stereo vision geometry for depth measurement is illustrated in Figure 3. The stereo depth camera used in this study, RealSense 435i, uses two infrared sensors to measure the depth value.

The procedure for measuring the distance between two objects by utilizing the measured depth values is as follows. (1) The full width of the camera Field of View (FOV) at a specific distance from the camera to the object is applied at an absolute value. (2) The total horizontal length of the camera FOV is calculated from the position of the object in front of the objects to be measured in proportion to the previously applied value. (3) Obtain the depth value of the object in front through Object Detection (Set to X). (4) Set the difference between depth values between objects to Y. The distance R between objects is calculated by following equation [R^2 = X^2 + Y^2].

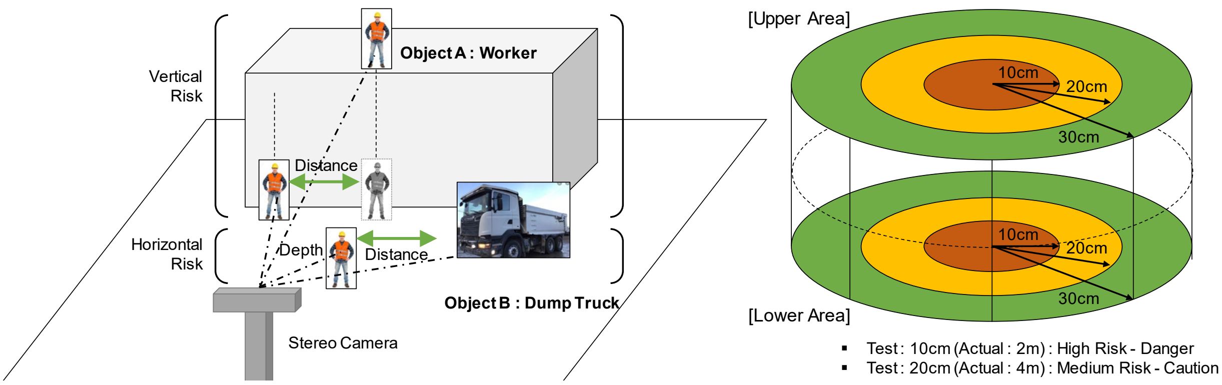

Figure 2. Lab Test Conditions

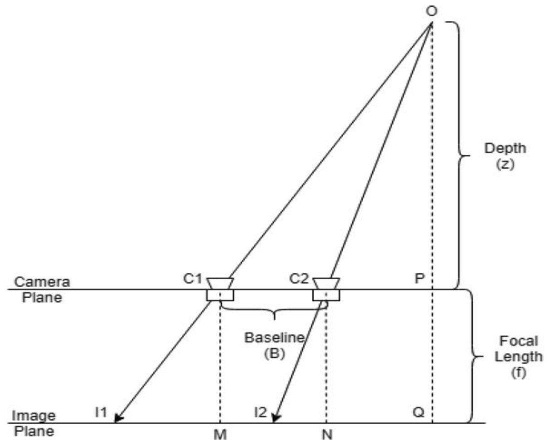

Figure 3. The Stereo Vision Geometry for Depth Measurement [14]

4.2 Lab Test Results

Lab test was performed on the laptop system with specifications as follows; Intel Core i7 8700K, 16 GB RAM, GTX1080TI VGA, Intel RealSense435i, Ubuntu 18.04, and Tensorflow 2.1.0.

As shown in Figure 4, the horizontal risk level (danger, caution, and safe) was determined in real time according to the set risk area as the worker approaches the dump truck.

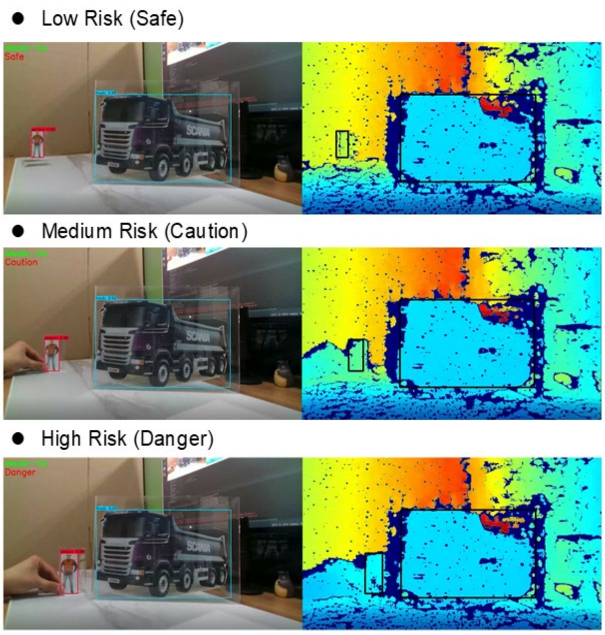

Figure 4. Example of Real-time Horizontal Risk Checking

As shown in Figure 5, motion detection determines the motion by comparing pixel changes between the current frame and a certain number of previous frames.

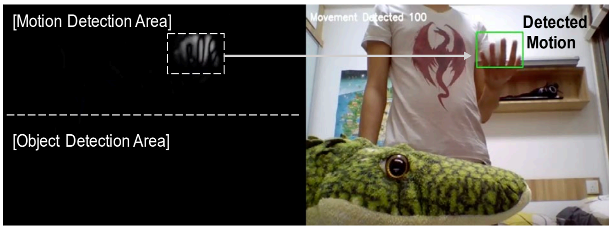

Figure 5. Example case of a motion detection

As described in the system process, when an object that moves at the top of the video frame is detected, the object detection algorithm is applied to determine the risk based on the distance between objects. In this test, the position of the lower worker was fixed and the upper worker was approached toward the lower worker. The difference between the X-axis value of the left pixel of the upper worker and the right pixel X-axis value of the lower worker was calculated as the horizontal distance.

As shown in Figure 6, the vertical risk level (danger, caution, and safety) was determined in real time according to the set risk area as the upper worker approaches the lower worker.

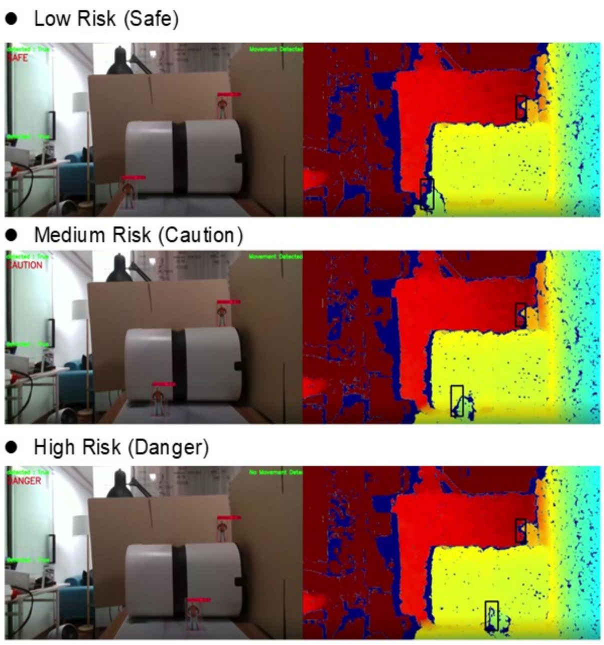

Figure 6. Example of Real-time Vertical Risk Checking

Table 1 and 2 show the results of vertical and horizontal risk level measurement. Tests were performed by measuring the depth and distance values by placing objects on the basis of the risk level (10 cm, 20 cm, 30 cm).

In the horizontal hazard test, when the distance between the worker and the dump truck was placed at 30cm, the measured distance was 31.6cm, with an error of 5.3%, but when placed at 10cm, the measured distance was 8.81cm, with a slightly larger error of 11.9%. The size of the 1/20 scale model used for the test was small, so close-up shots were taken, which caused the infrared sensors to be unstable.

Table 1. Results of Horizontal Risk

Actual Distance (Risk Level)	Depth Measurement		Distance between Objects	Error
	Worker	Dump truck		
10cm (Danger)	30cm	25cm	8.81cm	11.9%
20cm (Caution)	40cm	25cm	18.5cm	7.5%
30cm (Safe)	55cm	25cm	31.6cm	5.3%

In the horizontal hazard test, when the distance between the workers was placed at 30cm, the measured distance was 29.7cm, with an error of 1%, but when placed at 10cm, the measured distance was 9.8cm, with an error of 2%. In the case of vertical hazard determination, the error was relatively low because the distance was measured using the RGB image-based pixel information, not the infrared sensor-based depth measurement method.

Table 2. Results of Vertical Risk

Actual Distance (Risk Level)	Depth Measurement		Distance between Objects	Error
	Worker (Lower Area)	Worker (Upper Area)		
10cm (Danger)	45cm	57.4cm	9.8cm	2%
20cm (Caution)	45cm	52.8cm	20.5cm	2.5%
30cm (Safe)	45cm	50cm	29.7cm	1%

5 Conclusions

This study proposed horizontal and vertical risks detecting system around workers by measuring distance between objects using stereo vision technology, with the goal of preventing accidents caused by unrecognized hazards such as material drop during lifting and collision with construction vehicle. In this paper, motion detection and object detection are designed to operate sequentially in one video information to minimize computing resources.

The proposed system is tested in the lab environment. Test results showed that objects moving horizontally and vertically were detected in real time, and hazard levels were also determined in real time. However, the accuracy of the measured distance was low due to the size of the scale model utilized in the lap test and the narrow distance between the infrared lenses of the camera used.

In the future research, in order to increase the accuracy and applicable range of distance measurement between objects, we intend to test the measurable distance and viewing angle by using various stereo depth cameras on the actual field.

Acknowledgement

This work is supported by the Korea Agency for Infrastructure Technology Advancement(KAIA) grant funded by the Ministry of Land, Infrastructure and Transport (Grant 20SMIP-A158708-01).

References

[1] Wang, J.; Asce, S.M.; Razavi, S.N.; Asce, M. Low False Alarm Rate Model for Unsafe-Proximity Detection in Construction.

[2] Kim, K.; Kim, H.; Kim, H. Image-based construction hazard avoidance system using augmented reality in wearable device. *Autom. Constr.* **2017**, *83*, 390–403.

[3] Lee, H.-S.; Lee, K.-P.; Park, M.; Baek, Y.; Lee, S. RFID-Based Real-Time Locating System for Construction Safety Management. *J. Comput. Civ. Eng.* **2012**, *26*, 366–377.

[4] Rock, L. BIM and RFID integration : A pilot study. *Adv. Integr. Constr. Educ. Res. Pract.* **2002**, 570–578.

[5] Lu, Q.; Lee, S. Image-Based Technologies for Constructing As-Is Building Information Models for Existing Buildings. *J. Comput. Civ. Eng.* **2017**, *31*, 04017005.

[6] Ju, Y.; Kim, C.; Kim, H. RFID and CCTV-Based Material Delivery Monitoring for Cable-Stayed Bridge Construction. **2012**.

[7] Brilakis, I.; Park, M.-W.; Jog, G. Automated vision tracking of project related entities. *Adv. Eng. Informatics* **2011**, *25*, 713–724.

[8] Teizer, J.; Cheng, T. Proximity hazard indicator for workers-on-foot near miss interactions with construction equipment and geo-referenced hazard areas. *Autom. Constr.* **2015**, *60*, 58–73.

[9] Teizer, J.; Vela, P.A. Personnel tracking on

construction sites using video cameras. *Adv. Eng. Informatics* **2009**, *23*, 452–462.

[10] Park, M.W.; Makhmalbaf, A.; Brilakis, I. Comparative study of vision tracking methods for tracking of construction site resources. *Autom. Constr.* **2011**, *20*, 905–915.

[11] Seo, J.; Han, S.; Lee, S.; Kim, H. Computer vision techniques for construction safety and health monitoring. *Adv. Eng. Informatics* **2015**, *29*, 239–251.

[12] Mchugh, E.L. VIDEO MOTION DETECTION FOR REAL-TIME HAZARD WARNINGS IN SURFACE MINES Edward L. McHugh Spokane Research Laboratory, National Institute for Occupational Safety and Health, Spokane, Washington, USA.

[13] Saxena, A.; Schulte, J.; Ng, A.Y. *Depth Estimation using Monocular and Stereo Cues*;

[14] Praveen, S. Efficient Depth Estimation Using Sparse Stereo-Vision with Other Perception Techniques. In *Coding Theory*; IntechOpen, 2020.