

Entity Matching across Stereo Cameras for Tracking Construction Workers

Yong-Joo Lee^a, Man-Woo Park^b, Ioannis Brilakis^c

^{a, b} Department of Civil and Environmental Engineering, Myongji University, Republic of Korea

^c Department of Engineering, University of Cambridge, United Kingdom

E-mail: leetoday@mju.ac.kr, mwpark@mju.ac.kr, ib340@cam.ac.uk

Abstract

While GPS and RFID are dominantly used for tracking construction equipment and materials, 3D vision-based tracking that employs stereo-vision system has been also proposed as an alternative to efficiently track construction resources operating in a congested site. The 3D vision-based tracking requires an entity to be tracked in both cameras of the stereo-vision system so that two of 2D pixel coordinates are generated for each entity. Its 3D coordinate can be calculated by triangulating two projection rays that go through the 2D pixel coordinates. In order to track multiple entities of the same type simultaneously, the triangulation requires an additional process to match entities across the two camera views. This paper proposes an efficient method of entity matching for tracking construction workers. The method is mainly based on epipolar geometry which represents the geometric relation between two camera views. Two additional strategies – distance ratio thresholding and frame grouping – are introduced to reject false matchings and to ultimately increase precision of the 3D tracking. Experiment results demonstrate its suitability for automated entity matching for 3D vision-based tracking of construction workers.

Keywords –

Construction worker; Stereo Vision; Tracking; Detection; Entity Matching

1 Introduction

Any given construction project is composed of a series of processes through the organic and diversified interactions of manpower, equipment and materials. These processes in turn are composed of complicated tasks that are difficult to anticipate. Therefore, it is important to manage the schedules and extent of progress by minimizing uncertainties that can occur in construction sites in order to successfully execute construction projects. Uncertainties ultimately arise from

lack of information. With a greater amount of and more diversified information, site managers and decision makers are able to make decisions that are more effective and safer. Data on the location and conditions of the construction equipment, manpower and materials, etc. can be used as important data for the assessment of the status of the construction sites. Research on the utilization of such data have been conducted for the purposes of management of on-site safety and productivity, along with the active execution of research on the position tracking system for each of the types of entities.

Global Positioning System (GPS) is one of the most representative position tracking systems. It is highly reliable since it is possible to secure extremely high accuracy and can be used in all regions provided it is in outdoor settings depending on the composition of the system, and there are many models of the latest heavy machineries with built-in GPS. Recently, Radio Frequency Identification (RFID) [1], Ultra Wide Band (UWB) [2] and Bluetooth [3] formats of position tracking technologies have been introduced to the market. However, the majority of the position tracking methods including GPS need to attach a tag to each entity to be tracked. Accordingly, due to the need for the management of the system such as reading the tag or replacing of the aged tag at prescribed intervals, the cost of position tracking increases rapidly if the number of entities to be tracked increases substantially. However, the image-based position tracking system only uses a camera and processor for the image processing without the need for a separate tag. Due to such advantage, research on the image-based position tracking system has been conducted actively. Research on real-time monitoring of the construction sites by using a 3D imaging sensor and managing the materials by checking the location and conditions of the construction equipment and manpower through the image information and by recognizing the construction materials have been executed [4][5][6]. Image-based tracking system is being recognized as a technology capable of efficiently providing key data for the assessment of the movement

paths and patterns of the equipment, manpower and materials in the construction sites.

Majority of the image-based process management methods introduced above were limited to the obtaining of 2D pixel coordinates by assessing the movement and transformation of the entities by using monocular camera. Stereo image is necessary to compute the 3D coordinates in the actual space of particular entity from the 2D images. Park et al. proposed the method of acquiring the actual 3D position coordinate using a stereo vision system [5]. They showed that it is possible to estimate the 3D coordinates of a given entity through the triangulation among the entity photographed and its projected pixel coordinates on two cameras [5]. However, the validation was limited to tracking of a single entity. An additional separate method is necessary in order to track multiple entities. Acquiring 2D coordinates of the entities in each of the images should be followed by the matching of the same entities in the left and right images, which is not needed when tracking a single entity since the matching is straightforward. This paper proposes an efficient method of entity matching for tracking construction workers. The method is mainly based on epipolar geometry which represents the geometric relation between two camera views, and the preliminary experiments showed its suitability for the stereo vision based 3D tracking.

2 Background

GPS and RFID have attracted a lot of attention from construction industry as they can be used for tracking the positions of construction entities such as workers, equipment, materials. As an alternative to GPS and RFID, the vision-based tracking system has also been investigated because of its several potential advantages. For example, it does not require to tag every entity to be tracked, and it can potentially track any entities that appear in the camera view [7]. Continuous research efforts have been made on the image processing and computer vision algorithms to achieve the aforementioned potential advantages. Object detection and object tracking are the main algorithms required to locate the positions of interested objects in the 2D pixel coordinate system.

Chi and Caldas proposed the method of detecting the construction entities by using the background subtraction and the machine learning classifiers [8]. The construction entities were categorized by using the area size, aspect ratio and occupancy percentage, etc. separated by the background subtraction in the artificial neural network. Rezaadeh and McCabe proposed the method that can be used in the progress management through the tracking of the dump truck in the construction sites [9]. The method applied the Haar-like features and the HOG (Histograms

of Oriented Gradients) features to detect the dump trucks. Training data are generated by using the images of dump trucks viewed from various directions for detection. They also presented a part-based method to detect articulated shapes of machineries. Focusing on detecting construction workers, Park and Brilakis proposed the method to detect people wearing safety vests of fluorescent colours, which is composed of three steps - background subtraction, HOG + SVM (Support Vector Machine), and colour histogram + k-NN (k-nearest neighbour) [10]. Similarly, Memarzadeh et al. integrated the HOG and hue-saturation-based color histogram features into a single feature [11]. The feature is trained with SVM, and applied to workers, excavators, and dump trucks.

Whereas the methods described above were focused on detecting construction entities, several studies have been conducted on combining detection and tracking methods for continuous position tracking throughout video frames. Rezaadeh Azar et al. [12] employed a HOG based detection method and the KLT (Kanade-Lucas-Tomasi) tracker. The method is capable of tracking excavators and dump trucks even when they are partially occluded. Similarly, Park and Brilakis combined the function of detection and tracking algorithms for locating construction equipment [13] and workers.[14] Their methods stably handled total and partial occlusions by automatically identifying the events of occlusion based on the absence of detection results.

Park et al. [5] proposed the method of acquiring the 3D position coordinates by using a template-based 2D tracker [15] and applying the stereo vision technology. A fixed stereo vision system is used, and the geometric relation between two cameras were estimated by executing the camera calibration process through the use of SIFT (Scale-Invariant Feature Transform) and MAPSAC (Maximum A Posteriori Sample Consensus). The 2D pixel coordinates of the targeted entity is obtained by using a template-based tracking method in the left and right video frames. At each frame, the two 2D pixel coordinates of the entity extracted from the left and right video frames are fed into the triangulation process which calculates the 3D coordinate based on the geometric relation of the two cameras. The method was tested on tracking of a van, a worker, and a steel plate held by a worker. Though the tests showed reasonably high accuracy, each test involved a single target entity. The method is not applicable to tracking more than two entities since it is not able to distinguish the matching pair across the two views. An additional process is required to match the 2D detection/tracking results between the left and right camera views. Therefore, this paper proposes an efficient method for finding the matching pairs across the two views to enable the 3D tracking of multiple construction workers.

3 Eipolar Geometry based Entity Matching for Multiple Workers

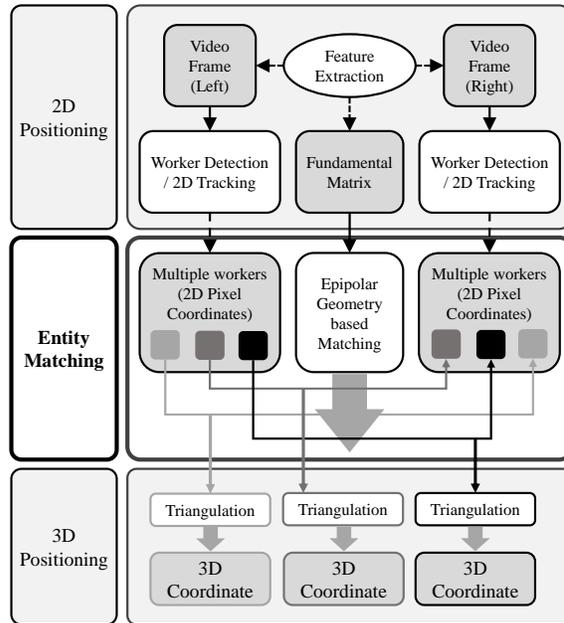


Figure 1. Framework for 3D tracking of multiple construction workers

Figure 1 illustrates the framework for the 3D tracking of multiple workers using a stereo vision system. First, the combination of the detection and 2D tracking methods is executed by receiving input from the two camera views. For this purpose, the method proposed by Park and Brilakis [14] is adopted. This step provides 2D pixel coordinates of the workers in each camera view. If three workers are detected and tracked in both camera views, three of 2D pixel coordinates will be provided continuously throughout the video frames from both views, respectively. Calculating the 3D coordinate of each worker is processed through triangulation which requires a pair of its 2D coordinates – one from the left frame and the other from the right frame. Therefore, the 3D tracking of a worker can be initiated only when the worker entity is found from both views and matched across the two views. As illustrated in Figure 1, the three worker entities found from both views need to be matched each other to determine three pairs of the 2D coordinates, each of which is corresponding to the same worker entity. It should be noted that the matching process is required only once to initiate the 3D tracking and no further matching process afterwards. This paper proposes an efficient method of the entity matching that mainly uses the epipolar geometry information.

3.1 Epipolar geometry based entity matching

Figure 2 shows brief summary of the matching principles based on the epipolar geometry. The epipolar geometry illustrates the geometric correlation between two cameras in terms of the relative position of the camera origins (O_L and O_R) and their view angle difference. For a worker positioned at x in actual 3D space, an epipolar plane is determined as the plane containing three points x , x_L , and x_R , where x_L and x_R are the 3D-to-2D projections of x onto the left and right image planes, respectively. The left projection line passing through x_L is projected onto the right image plane as l_R which is referred to the epipolar line of x_L . The epipolar line connects x_R and e_R which are the projections of x and O_L onto the right image plane, respectively. This geometric configuration shows that the 2D projection of the same entity onto the right image plane is located on the epipolar line l_R . Therefore, there must be image coordinates, among the pixel coordinates that epipolar line passes through, onto which the targeted entity has been projected. Accordingly, the primary matching principle is established that the matching pair of a 2D position tracked in the left view should be found from its corresponding epipolar line in the right. Complying with this principle, the matching pair of a worker tracked in the left view can be determined as the one appears closest to its epipolar line in the right view. This principle also works in the opposite way. In other words, the matching pair of a worker tracked in the right view can be determined as the one appears closest to its epipolar line in the left view.

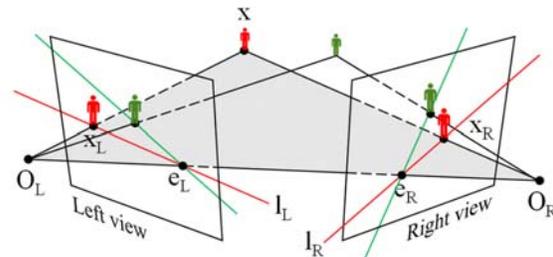


Figure 2. Relation of global 3D points and epipolar lines on each images

Figure 3 shows an example of the matching process. The figure illustrates the process of finding the matching pair of the worker entity which is represented as a bounding box (A) in the right view. In the figure, there are three candidate worker entities detected and tracked in the left view ((a), (b), (c)). The red line drawn in the bottom-left image is the epipolar line that corresponds to the centroid of the bounding box (A). In order to find the matching entity from the left view, the distance from the

bounding box centroid to the epipolar line is calculated for each candidate. The candidate (c) is determined as the matching entity since its distance d_c is the smallest.

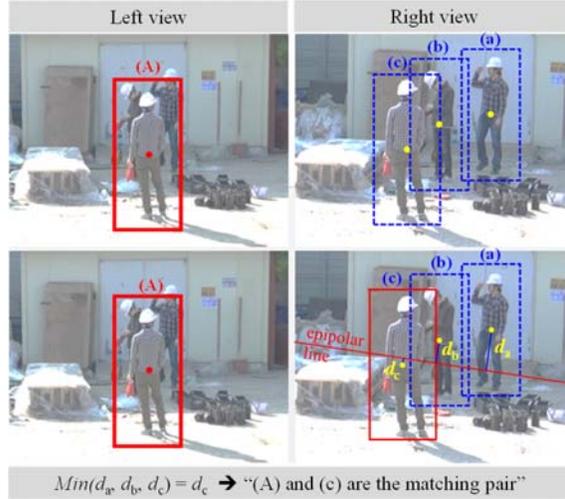


Figure 3. Matching process based on the distance from entity centroids to the epipolar line

The epipolar geometry is generally represented by the 3-by-3 fundamental matrix. The fundamental matrix between the two cameras can be computed by extracting the feature points independently from the images taken by each of the cameras [16]. Since the calculation of the fundamental matrix by using an accurate pair of feature points can accomplish even more accurate entity matching, it is important to use an algorithm that can extract the feature points without being affected by the changes in scale, rotation and distortion of the images. There are diverse methods for the extraction of such feature points. FAST (Feature from Accelerated Segment Test) [17], SURF (Speeded-Up Robust Features) [18] and SIFT [19] are known to be generally outstanding in the area of computer vision are compared. FAST is generally used for mobile applications which require real time processes. In this research, the calculation of the fundamental matrix is processed only once after the camera system has been fixed. Therefore, the real time process is not an important issue. SURF and SIFT would be good options for this research because of their ability to provide a large number of matching points with high accuracy. The number of matching points becomes more important when the image distortion is not calibrated well enough. By using the fundamental matrix computed with SURF or SIFT, the epipolar line corresponding to the centroid of the detected/tracked entities is generated.

3.1.1 Distance ratio thresholding

Finding the tracked entity located closest to the epipolar line does not always give the correct matching. There are several problems to be resolved to implement the epipolar geometry based entity matching. The first problem is the case when the corresponding entity is not detected or tracked in the opposite view. For example, if the entity labelled as (c) is not detected or tracked in the left view, then the matching entity of (A) would be determined as (b) which is the second closest. However, the candidate (b) is not the correct matching and locates much farther than the actual matching entity. In order to remove this type of false matching and minimize the effect of the false results from the detection and 2D tracking methods, this paper introduces the distance threshold. The distance threshold prevents the matching with the entity at a distance too far in this process.

The second problem is the case when two or more entities may be projected onto the single epipolar line. This case can happen if multiple entities are actually on the same epipolar plane. It can also happen even when entities are on the different epipolar plane because of the errors of the detection and 2D tracking methods. Accordingly, it is necessary to introduce an additional procedure to solve the problem. For this purpose, distance ratio threshold was additionally applied. The distance ratio thresholding minimizes the false matchings and maximize precision. The distance ratio (see Equation (1) and Figure 4) illustrates the ratio of the distances of the entities that are the closest and the second closest from an epipolar line. Under the presumption that it is not possible to determine which of the entities is the correct pair if this ratio is too low, the closest entity will be matched to the epipolar line only when the distance ratio is higher than the threshold.

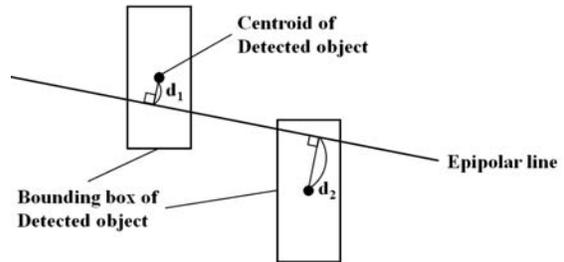


Figure 4. Distance between an epipolar line and the centroids of tracked objects

$$\text{Distance ratio} = \frac{d_2}{d_1} \quad (1)$$

The entity matching process is executed by using these two thresholds. Following is the summary of this process for finding the matching pair of each entity

tracked in the left camera view. The matching pair is searched in the right camera view.

1. If the pixel distance between the epipolar line and the centroid of the entity in the right camera view is smaller than the distance threshold (T_1), it is deemed to have been matched.
2. If there are more than two entities on a single epipolar line,
 - The closest one is determined to be the matching entity only if the distance ratio is higher than the distance ratio threshold (T_2).
 - If the distance ratio is lower than the distance ratio threshold (T_2), both entities will be excluded from matching.

The above process is repeated in the opposite way – finding the matching pair of the tracked entities in the right camera view. The entities that produced the same results in both ways are matched. While this thresholding process is helpful for reducing false matchings, it also cuts true matching results. In other words, it enhances the matching precision, but degrades the matching recall. It should be noted that the precision is more important than the recall in this research since the matching process is required only once at the beginning of the tracking for each entity. The critical factor is not providing the matching results every frame, but reducing false matchings.

3.1.2 Frame Grouping

Due to the errors that can occur in the worker detection and 2D tracking processes, there are frequent situations in which accurate matching cannot be performed even if various thresholds are mobilized. To enhance the matching precision even further, this paper introduced the frame grouping strategy that determines the matching results for a certain interval, instead of every frame. It integrates the results of several frames in the format of voting in accordance with the set frequency.

Figure 5 illustrates the frame grouping process for searching the matching result of the worker on the farthest left in the image, which is labelled as '1'. From top to bottom, the figure are the results of the individual matchings for the 348th, 349th, and 350th frames, respectively. If the frame grouping is not applied, the method would generate two correctly matched results (#348 and #349) and one wrongfully matched result (#350). However, if frame grouping is applied, the three frames are grouped together and the one false result would be neglected due to the majority votes of the correct results. In other words, instead of providing two correct results and one false results, it generates one true matching result. As such, it is possible to anticipate more

accurate results if frame grouping is applied. Although results are obtained more intermittently if the number of the frames grouped together increases, its effect on precision will manifest to greater extent.

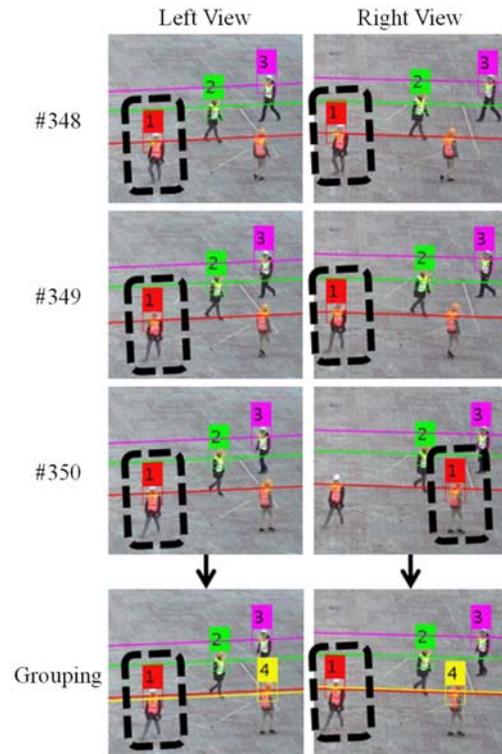


Figure 5. Frame grouping process for the three consecutive frames

4 Experimental Results

For the experimental tests, videos were taken of the 4 persons wearing hardhats and safety vests. These subjects were allowed to walk freely within the prescribed area. The original images in MTS format were edited to the length of approximately 40 seconds and were encoded in MP4 format while maintaining the same number of frames. The propose framework was processed for 10 fps and the total number of the processed frames was 398.

4.1 Fundamental Matrix Calculation

As a first step to implement the proposed matching framework, the fundamental matrix is calculated using the feature point matching results. In this research, no intrinsic calibration is performed, so camera distortion is not corrected. Therefore, a large number of feature

matchings are required to obtain more accurate fundamental matrix. A simple experiment is performed to compare the SIFT and the SURF in terms of the number of correct matching points. For both method, RANSAC (RANDOM Sample Consensus) [20] is used to remove outliers and to compute the fundamental matrix as accurately as possible. The test showed that the SIFT and the SURF provide 954 and 224 matching pairs, respectively. Accordingly, the SIFT is used for the fundamental matrix calculation. Since the entity matching method proposed in this research uses fixed cameras, it only uses a single fundamental matrix that was computed based on the first frames of the left and the right camera views.

4.2 Epipolar Geometry based Entity Matching with Thresholding

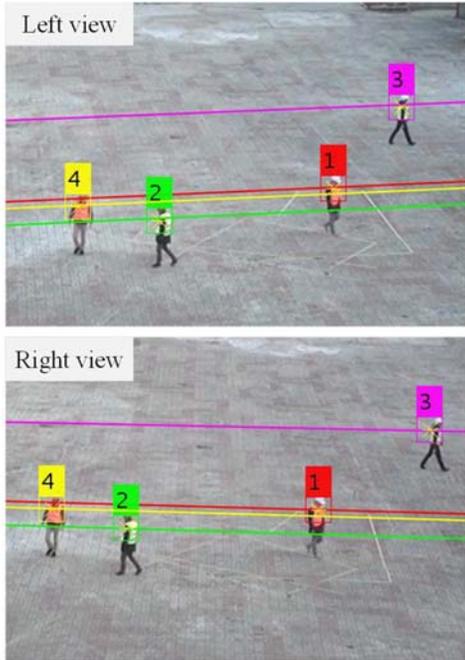


Figure 6. Correct matching results from the 257th frame ($T_1 = 10$, $T_2 = 2.5$)

Firstly, the experiments were conducted without using frame grouping to determine the appropriate value of the distance ratio threshold (T_2). Figure 6 and Figure 7 illustrate some example results. Entities determined to be the same worker are indicated with the same coloured bounding box and ID number. Lines drawn with the same colour used for the bounding box indicate the epipolar line of the corresponding entity. The epipolar line of the entity tracked on the left is drawn on the right video frame

and vice versa. Figure 6 illustrates the correctly matched results while the Figure 7 illustrates the results in which the entities labelled as '1' and '4' are falsely matched. If multiple entities are confirmed on the same epipolar line, it may cause a false matching as illustrated.

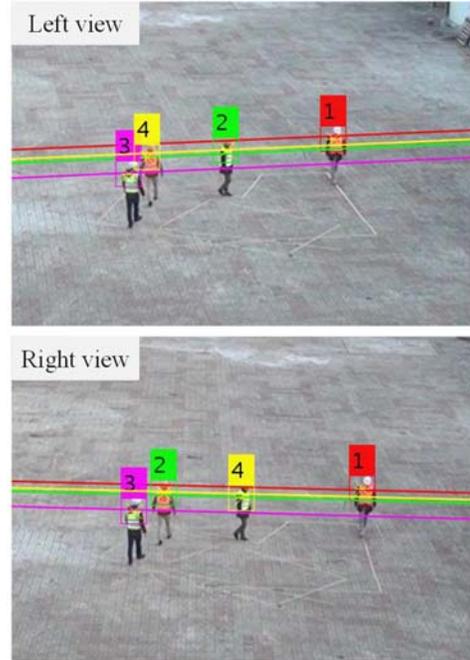


Figure 7. False matching results from the 35th frame (Threshold: distance 10, distance ratio 2.5)

In order to investigate effect of distance ratio thresholding, the recall and the precision are calculated for various threshold values. The precision and the recall are defined as follows.

- Precision = $TP / (TP + FP)$
- Recall = $TP / (TP + FN)$
- TP (True Positive) = # of the retrieved matchings which are correct
- FP (False Positive) = # of the retrieved matchings which are incorrect
- FN (False Negative) = # of the matchings which are not retrieved

Table 1. Results of entity matching according to the distance ratio threshold

Distance ratio threshold, (T_2)	Recall (%)	Precision (%)
Not Applied (0.0)	91.21	96.49
Applied (2.5)	80.72	98.10
Applied (4.0)	71.40	98.65

The experiment was conducted by changing the distance ratio threshold from 1.1 to 4.0 at the interval of 0.1 in order to determine the appropriate threshold value T_2 . Throughout the experiments, the distance threshold T_1 is maintained equal to 10. Table 1 and Figure 8 illustrate the summary of the results. When the thresholding is not applied ($T_2 = 0$), the proposed method scored 96.49% precision, which is increased to 98.10% by applying $T_2 = 2.5$. The precision was found to be higher when the higher threshold was applied although the recall falls to 70%. As mentioned in the previous section, the entity matching method introduced in this research places greater importance on the precision rather than on the recall. Therefore, it was decided to use distance ratio threshold of 4.0 because no further precision enhancement is observed with $T_2 > 4.0$.

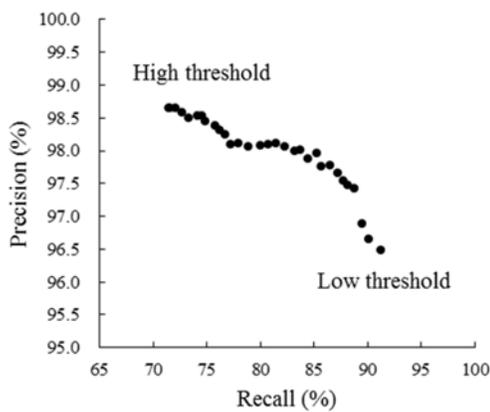


Figure 8. Changes in the recall-precision according to the distance ratio threshold set

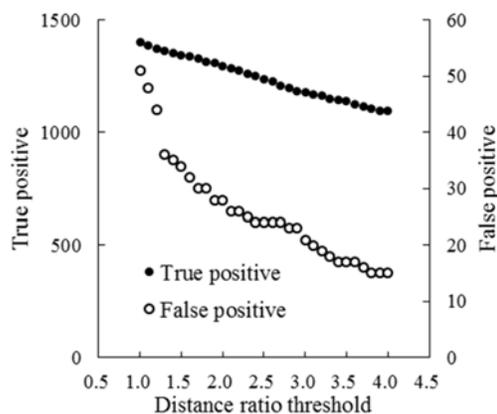


Figure 9. Changes in TP and FP according to the distance ratio threshold set

Figure 9 illustrates the numbers of the true positives

and false positives in accordance with T_2 . There is rapid improvement of performances in the range of 1.0~1.5 with gradual improvement of performances thereafter. Although the number of entities that are correctly detected decreases with the increase in T_2 , the overall precision actually improves as the false matchings are reduced far more than the decrease in the number of the true matchings.

4.2.1 Epipolar Geometry based Entity Matching with Frame Grouping

Table 2 compares the entity matching results with and without using the frame grouping. Both experiments were conducted under the condition of $T_2 = 4.0$. When applying the frame grouping, the number of the frames to be grouped was increased from 3 to 13 at the interval of 2 frames. Only odd numbers were used since the even numbers can bring the score to a tie. Comparisons were made by using the results at the frequency 13, which has the highest precision among all the groupings. Slightly higher precision is obtained by applying the frame grouping. It exhibits substantially lower recall value which is a predictable result because the frame grouping provides a matching result every 13 frames. According to the results, it is fair to expect the proposed method with frame grouping to provide the matching pair of each entity in a second with 99.17% accuracy.

Table 2. Entity matching results with/without frame grouping (the number of frames in a group = 13)

Frame Grouping	Precision (%)	Correct matchings	Wrong matchings
Not Applied	98.65	1096	15
Applied	99.17	119	1

5 Conclusion

Research efforts have been made on the vision-based tracking for positioning construction entities in construction sites. The 2D localization in a single camera view through detection and 2D tracking methods has been investigated in a series of research works. Furthermore, it was reported that 3D locations of the entities can be obtained by employing stereo vision system. This paper proposes an efficient method of entity matching which is essential for tracking multiple entities using the stereo vision. It was confirmed that the matching between the same entities in the stereo videos is possible by applying the epipolar geometry onto the results of the worker detection and 2D tracking. In order to maximize the precision which is considered as the most critical factor of the matching performance in this

research, the distance ratio thresholding and frame grouping are introduced. The preliminary results showed that the distance ratio thresholding could greatly reduce the false matchings which was reflected in the higher precision results. Also, the frame grouping could enhance the precision even further achieving the precision over 99%. The preliminary results showed the potential of the proposed method that can be fitted in the 3D tracking framework to finally enable acquiring 3D positions of multiple workers simultaneously.

Acknowledgement

This material is based upon work supported by the Grant No. 14CTAP-C078828-01 from Infrastructure and Transportation Technology Promotion Research Program funded by the Ministry of Land, Infrastructure and Transport of Korean government. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Ministry of Land, Infrastructure and Transport of Korean government.

References

- [1] Song J., Caldas C., Ergen E., Haas C. and Akinci B. Field trials of RFID technology for tracking pre-fabricated pipe spools. In Proceedings of the 21th International Symposium on Automation & Robotics in Construction, 2004.
- [2] Teizer J., Lao D. and Sofer M. Rapid automated monitoring of construction site activities using ultra-wideband. In Proceedings of the 24th International Symposium on Automation & Robotics in Construction, pages 23-28, 2007.
- [3] Hallberg J., Nilsson M. and Synnes K. Positioning with Bluetooth. In Proceedings of the 10th International Conference on Telecommunications 2003, pages 954-958, 2003.
- [4] Teizer J., Caldas C. and Haas C. Real-Time Three-Dimensional Occupancy Grid Modeling for the Detection and Tracking of Construction Resources. *Journal of Construction Engineering and Management*, 133(11):880-888, 2007.
- [5] Park M., Koch C. and Brilakis I. Three-Dimensional Tracking of Construction Resources Using an On-Site Camera System. *Journal of Computing in Civil Engineering*, 26(4):541-549, 2012.
- [6] Park M., Makhmalbaf A. and Brilakis I. Comparative study of vision tracking methods for tracking of construction site resources. *Automation in Construction*, 20(7): 905-915, 2011.
- [7] Brilakis I., Park M. and Jog G. Automated vision tracking of project related entities. *Advanced Engineering Informatics*, 25(4):713-724, 2011.
- [8] Chi S. and Caldas C. Automated object identification using optical video cameras on construction sites. *Computer-Aided Civil and Infrastructure Engineering*, 26(5):398-380, 2010.
- [9] Rezaazadeh E. and McCabe B. Automated visual recognition of dump trucks in construction videos. *Journal of Computing in Civil Engineering*, 26(6): 769-781, 2012.
- [10] Park M. and Brilakis I. Construction worker detection in video frames for initializing vision trackers. *Automation in Construction*, 28:15-25, 2012.
- [11] Memarzadeh M., Golparvar-Fard M., and Niebles J. C. Automated 2D detection of construction equipment and workers from site video streams using histograms of oriented gradients and colors. *Automation in Construction*, 32:24-37, 2013.
- [12] Rezaazadeh Azar E., Dickinson S., and McCabe B. Server-Customer Interaction Tracker: Computer Vision-Based System to Estimate Dirt-Loading Cycles. *Journal of Construction Engineering and Management*, 139(7):785-794, 2013.
- [13] Park M. and Brilakis I. Enhancement of Construction Equipment Detection in Video Frames by Combining with Tracking. In Proceedings of ICCCE 2012, pages 421-428, Clearwater Beach, Florida, United States, 2012.
- [14] Park M. and Brilakis I. Improved localization of construction workers in video frames by integrating detection and tracking. The CSCE International Construction Specialty Conference, Vancouver, Canada, 2015.
- [15] Ross D., Lim J., Lin R. and Yang M. Incremental learning for robust visual tracking. *International Journal of Computer. Vision*, 77(1):125-141, 2008.
- [16] Luong Q. and Faugeras O. The fundamental matrix: Theory, algorithms, and stability analysis. *International journal of computer vision*, 17(1):43-75, 1996.
- [17] Rosten E. and Drummond T. Machine learning for high-speed corner detection. In Proceedings of the 9th European Conference on Computer Vision, pages 430-443, Graz, Austria, 2006.
- [18] Bay H., Tuytelaars T. and Gool L. Speeded-Up Robust Features (SURF). *Computer Vision and Image Understanding*, 110(3):346-359, 2008.
- [19] Lowe D. Distinctive Image Features from Scale-Invariant Keypoints. *International journal of computer vision*, 60(2):91-110, 2004.
- [20] Martin F. and Robert B. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381-395, 1981.