www.itcon.org - Journal of Information Technology in Construction - ISSN 1874-4753

FEASIBILITY OF IN-PLANE ARTICULATION MONITORING OF EXCAVATOR ARM USING PLANAR MARKER TRACKING

SUBMITTED: July 2014
REVISED: December 2014
PUBLISHED: February 2015 at http://www.itcon.org/2015/15
EDITOR: Turk Ž.

Ehsan Rezazadeh Azar, Assistant Professor
Department of Civil Engineering, Lakehead University
Email: eazar@lakeheadu.ca

Chen Feng, PhD Candidate
Department of Civil and Environmental Engineering, University of Michigan
Email: cforrest@umich.edu

Vineet R. Kamat, Associate Professor
Department of Civil and Environmental Engineering, University of Michigan
Email: vkamat@umich.edu

SUMMARY: *Achieving accurate excavation profiles is a challenging task for excavator operators who typically use the directions of a grade-person to achieve design grades and levels. Allocation of an extra person is costly and requires constant communication between an operator and the grade-person, which reduces productivity. As a result, several machine control technology providers have developed angle sensor packages to monitor the pose of the excavator components in real-time. These systems, however, are expensive and their installation and calibration are costly and time consuming. This paper presents a generic and scalable computer-vision based framework for real-time pose estimation of an excavator's boom and dipper (stick) using low-cost markers installed on the side of the arms. The hardware components of this system are inexpensive and the setup process is quick and straightforward. The system demonstrated promising performance and has been shown to measure boom and dipper angles with an average uncertainty of less than 0.75° in real-time.*

KEYWORDS: *Machine Control, Pose Estimation, Computer Vision, Excavation, Marker-Based Recognition*

REFERENCE: *Ehsan Rezazadeh Azar, Chen Feng, Vineet R. Kamat (2015). Feasibility of In-Plane Articulation Monitoring of Excavator Arm Using Planar Marker Tracking. Journal of Information Technology in Construction (ITcon), Vol. 20, pg. 213-229, http://www.itcon.org/2015/15*

1 INTRODUCTION

Excavators are one of the most versatile pieces of construction equipment that can be used for bulk and precise excavation, loading dump trucks, trimming, trenching, and moving objects. In addition, there are several specialized add-ons, such as pneumatic hammers and roadheaders, which extend the fields of application for this type of machine. Operators are able to independently carry out many of these operations through self-visual observations. However, there are some precise operations in which the operators require external guidance to reach a clear-cut desired result. Precise excavation, trimming, and trenching are the most common processes that need to be carried out with high precision. For instance, the construction deviation for the slope of a sewage pipeline trench should be minimal as these systems use gravity to convey sewage. In the conventional approach, excavation companies assign an individual called the grade-person who continually measures the excavation profile and guides the excavator operator. At the end of the operation, surveyors need to check the excavated surfaces to calculate the excavated volumes and also to find deviations from the design profiles (Han et al., 2006). In addition to the design requirements, existence of underground utilities in brownfields complicates the excavation process and therefore further necessitates external guidance. Therefore, there has been a growing need to develop context-aware systems that can provide excavator operators with detailed information about the pose of the arm, and in particular the accurate location and orientation of the end-effector. Machine control systems are a relatively new solution to provide real-time spatial information for the machine operators and there are a variety of products for different earthmoving equipment which are mainly GPS-based. It is estimated that the application of these systems could improve the productivity of large operations by more than 20% (Han et al., 2006, Trimble Machine Control http://www.trimble.com/construction/heavy-civil/machine-control/).

A number of machine control technology companies, such as Trimble, Leica, and TOPCON, have developed relatively similar products to monitor the pose of the arm of excavators and backhoes. They have a range of solutions from one dimensional depth estimation to full 3D monitoring of the machine pose and its boom. The low-end solution in this product line includes a laser machine, laser receiver, and a display in which the laser catcher is installed on the dipper and notifies the operator if it has reached a certain depth. The next solution in their product line includes a set of tilt sensors attached on the boom, dipper, and bucket to estimate the angles of these components which are then transmitted wirelessly or via wires to a visual display in the cabin. These products cost from $9000 to $12000 and in addition, they require several hours of skilled-labor for calibration and installation which adds around $1000 to the total cost. Finally, it is difficult and costly to frequently attach and remove the sensor units due to security concerns. There are other sensor types that can be added to the system such as cabin sensors to measure the pitch and roll angles of the cabin. Electronic compasses are also introduced by some companies to measure the yaw angle of the boom although its performance in close proximity to several steel-made moving elements is uncertain.

Inclusion of an RTK GPS can provide the 3D location of the machine, but it would increase the entire cost of the system to more than $30000. The high-end solution in commercially-available product lines is a full 3D solution which uses two RTK GPS receivers to enhance 3D localization of the machine and also to measure the yaw angle of the equipment with higher accuracy. This GPS-based solution, however, is expensive and is priced around $55000 to $60000. It also suffers from the multipath problem associated with the GPS in areas with natural (such as trees) or human-made (such as urban canyons) obstacles (Meguro et al., 2009). High prices, difficulty of calibration, installation, and removal, and the loss of GPS signals have limited the application of these systems. Therefore, it would be worthwhile to investigate other emerging sensing technologies, which promise reliable and inexpensive performance. In particular, computer vision algorithms are among the fastest growing sensing technologies, which are rapidly being adopted in different industrial and service segments. This is mainly due to the fast improvements of computer vision algorithms, and emergence of low-cost digital cameras and processing platforms (such as smartphones, tablets, and laptops). Quick Responses code (QR code), product quality control, and automatic number plate recognition are instances of popular vision-based applications.

This article describes research on scientific questions regarding application of a vision-based method for excavator's boom monitoring, which investigates real-time performance, accuracy, and sensitivity to affecting factors including distance and angle of the optical axis. This research paper introduces a low-cost and user-friendly framework which uses special fiducial markers to estimate the angles of the boom and dipper of excavators and backhoes. This approach promises an economical and easy-to-use excavator monitoring system for the operations where the markers remain visible for the side camera, such as shallow trenching and trimming.

2 BACKGROUND STUDIES

Construction equipment monitoring has been extensively investigated at both macro and micro levels. At the macro level, users are interested in simultaneous localization of several machines in a fleet in real-time for productivity measurement, safety, and fleet management purposes. A variety of technologies including GPS (Navon and Shpatnisky, 2005, Navon et al., 2004), UWB (Teizer et al., 2008), and computer vision (Rezazadeh Azar et al., 2013, Memarzadeh et al., 2013, Rezazadeh Azar and McCabe, 2012a, Rezazadeh Azar and McCabe, 2012b, Gong and Caldas, 2011) have been applied to localize and track construction machines.

On the other hand, micro-level monitoring systems provide accurate pose of the end-effector for the machine operators to carry out processes quicker and with high precision. Accuracy of the measurements and the ability to provide data in real-time are essential in these systems. Common GPS-based systems are not able to provide the required precision and therefore a special technique called Real Time Kinematic (RTK) GPS, is used to enhance the precision of the measurements. RTK systems include a base station receiver and some mobile units. The base station transmits the phase of the carrier that it receives, and the mobile units use both their own phase measurements and the one received from the base station to calculate their 3D location.

The RTK systems are extensively used to guide excavation and leveling equipment. One or two mobile antennas are installed on the blade of bulldozers and graders to reach the design profiles faster while eliminating the need for the guidance of a grade-person. The end-effector of excavators, however, cannot be monitored with a RTK GPS, thus a system of tilt sensors are used to measure the angle of each part of the arm in its plane. One or two RTK GPS units can be added to provide 3D information of an operating machine. Optical sensors are another alternative for machine control which include retro-reflector or active optical sensor. Retro-reflectors are designed to reflect the light with minimal scattering and have been employed for precision localization, such as prism for total station, and robot guidance. Active optical sensors are attached to the target and are able to measure the location by emitting or receiving light, more specifically laser. This technique is commercially used for machine control purpose, such as grade control for backhoes, and excavators. This approach is a 1D machine control system in which a laser receiver is attached to the dipper of the machine to catch the laser rays emitted from a fixed laser machine and calculate the elevation of the dipper. This system, however, is not able to provide other pose data, including angles of the boom and accurate pose of the bucket.

In addition to commercial pursuits, some researchers have investigated detailed machine monitoring, including excavators. Specifically, micro-level equipment monitoring was extensively studied for design-integrated operation and for autonomous excavator. For instance, three electronic joint encoders were used to estimate the pose of the boom components in the excavator arm’s plane and two laser receivers were added to calculate the site coordinates for design-integrated excavation. The accuracy of excavation using this prototype was within 5.1 cm (Huang and Bernold, 1997, Bernold, 2002). In other research effort, two wireless cameras and a laser target were used for tele-operated pipe laying by an excavator (Lee et al., 2003, Bernold, 2007). Development of robotic excavator has been a popular topic for the last two decades and in some of the developed prototypes, the control unit of autonomous excavator perceives the pose of the arm elements using various sensing devices (Stentz et al., 1999, Chiang and Huang, 2004, Yamamoto et al., 2009). In addition to tele-operation and robotics, Talmaki and Kamat (2014) developed a system which uses the output of sensors to emulate the movement of an excavator in a virtual reality environment to estimate the minimum distance between the end-effector and buried utilities in excavator’s vicinity.

The main shortcomings of existing technologies for excavators are the high cost of sensing packages and the difficulty of calibrating and installing them. This research aims to close this gap by introducing a low-cost and user-friendly vision-based approach to monitor the movement of an excavator boom and dipper in real-time.

3 VISION-BASED RECOGNITION

The first challenge of a vision-based monitoring method is to recognize the components of an excavator arm, including its boom and dipper. Two recognition approaches, including model- and marker-based methods, could be used for this purpose. The algorithms in the first group seek for the target object using naturally-occurring features such as parts-, appearance-, and feature-based methods. The model-based recognition approach could get more challenging when the target varies considerably in appearance (Yang and Ramanan, 2011, Felzenszwalb et al., 2010). Human or pedestrian is the most investigated recognition target in the computer vision community due to its vast application in surveillance and traffic safety; however, the performance of state-of-the-art detectors are still far from ideal (Dollar et al., 2012).

An excavator is a highly deformable machine which makes it a difficult target for existing object recognition methods. A few studies have applied advanced recognition techniques to detect excavators and backhoes in a construction site, but they are only able to provide a bounding box and general pose of the machine which are still insufficient for accurate pose estimation. In addition, the levels of false positives and false negatives of those algorithms are too high for accurate machine control applications (Memarzadeh et al., 2013, Rezazadeh Azar and McCabe, 2012a, Chi and Caldas, 2011).

In the second approach, an artificial tag, called fiducial marker, is used to recognize a target, mostly in controllable environments. Fiducial markers are artificial landmarks installed in the field of view which usually consist of a set of easy-to-detect features to provide consistent detection and pose estimation results. The fiducial markers have a different goal than general 2D barcode systems, such as QR codes (Fig. 1.a), because they can also provide the camera position and orientation relative to a tag. Most of the fiducial markers use a set of black and white patterns to form simple geometric shapes such as straight lines, and sharp corners and edges.

Marker-based methods are intended to detect in low resolution, rotated, unevenly lit, or in the corner of an occluded image (Olson, 2011). One of the first and main application fields of fiducial markers has been Augmented Reality (AR) and therefore, popular libraries including ARToolKit (see Fig. 1.b) (Kato and Billinghurst, 1999) and ARTag (Fiala, 2005) have been developed to address this issue. These visual markers facilitate superimposing virtual objects. One of the most recent marker-based algorithms is Apriltag (Olson, 2011) which is recognized as a robust and computationally effective method (see Fig. 1.c). AprilTag algorithm firstly computes the gradient at every pixel, including their magnitudes and direction. Then, it clusters the pixels with similar gradient magnitude and directions using a graph-based algorithm, which then fits multiple line segments as the boundary of each cluster. Finally, different quadrilaterals are detected and used to decode a candidate tag/marker ID.

The Apriltag algorithm showed superior performance in experiments compared to other popular fiducial marker-based methods (Olson, 2011). This approach was also combined with some tracking methods to create a robust visual tracking tool in harsh construction environments (Feng and Kamat, 2013). Due to the robust performance and flexibly of this method under severe visual conditions, it was used in this research project to monitor the pose of an excavator arm. This method detects the boundaries of the markers with high precision which makes it a potential tool for precise pose estimation. Since the aim of the envisioned machine control system is to assist the excavator operator, it would be acceptable to add detachable visual markers to the machine which would be much easier compared to installation of current tilt sensors.

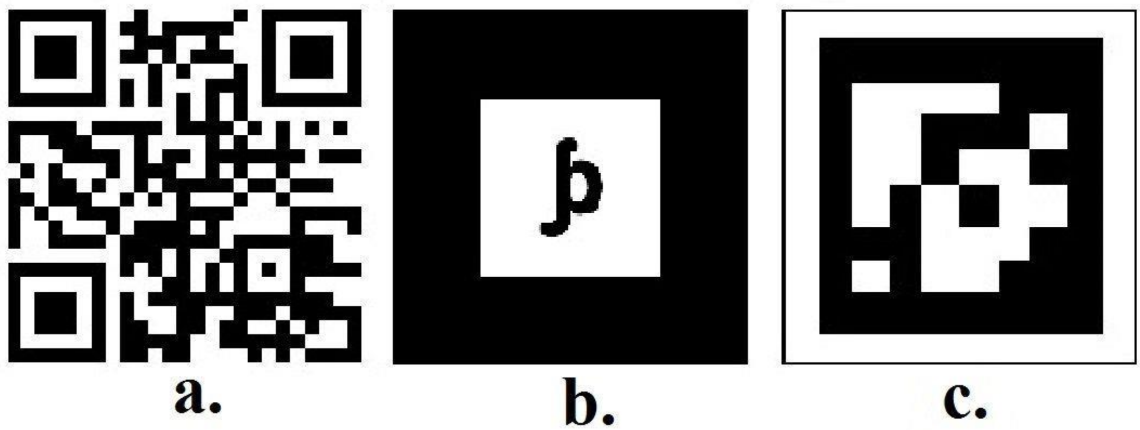

FIG. 1: Examples of (a) QR code (b) ARToolKit (Kato and Billinghurst, 1999) (c) AprilTag (Olson, 2011)

4 TECHNICAL APPROACH

Figure 2 represents the 2D kinematic chain of an excavator. Since the dimensions of each component of the arm is known, it is possible to estimate the 2D coordinates of all four points (A, B, C, and D) by knowing the coordinate of one point (such as A) and all three angles (θ_1, θ_2, θ_3). Therefore, the pose estimation can be maintained by having coordinates of one joint at the beginning and angle measurements during operation. One quick approach used by some commercial machine control systems is to touch the ground with the end-effector (point D) which allows the system to measure three angles and point D, and therefore calibrate itself (Prolec Machine Guidance, http://www.prolec.co.uk/2d_machine_guidance_.html). Then the system would be able to estimate the depth of the end-effector (point D) using the readings of the angles and kinematic calculation.

The main desired outcome is consistent and real-time measurement of the angles which are currently achieved using gravity-referenced sensors. This paper, however, introduces a framework to measure θ_1 and θ_2 using visual markers and a side camera. Angle measurement for θ_3 is not practical using this method as the bucket usually disappears from the field of view in excavation and trenching operations.

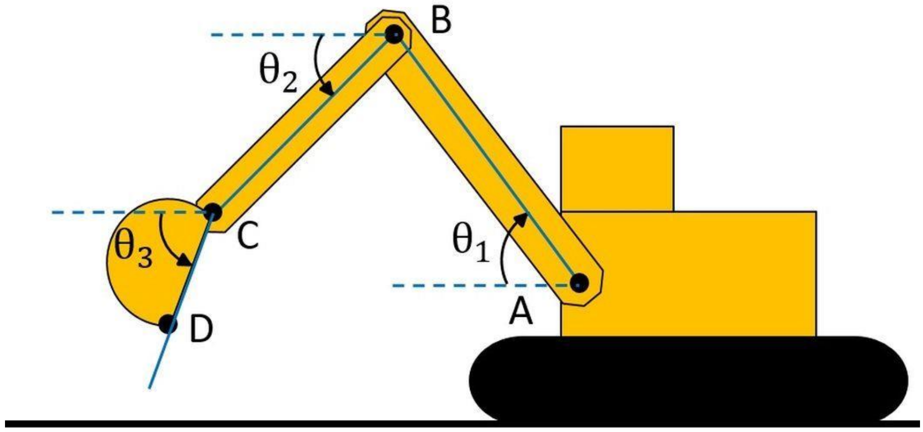

FIG. 2: Kinematic representation of excavator

The envisioned system uses low-cost fiducial markers (AprilTag), a regular digital camera, and a computer to estimate the pose of excavator arm in real-time. Markers' material and paint should not be reflective as they might be problematic in direct sunlight which could reduce marker detection rate. The Apriltag library was implemented in C++ (https://code.google.com/p/cv2cg/) which was the version used in this research project to develop the pose estimation framework. This system processes the video sequences to detect the tags in a form of two-dimensional quadrilaterals. The processing platform (could be a laptop, tablet, or smartphone) analyses all received video frames to detect markers. Since each tag has a unique identification number, different tags are installed on the boom and the dipper of the excavator to estimate the angle of those parts.

In this approach, two parallel markers were used to measure the angles of the boom and the dipper of a backhoe. Fig. 3 illustrates how the markers attached to the boom as well as the angle created as a result of marker recognition. The Apriltag algorithm estimates the boundary of the black square with high accuracy. If the system finds both parallel fiducial labels, the framework connects corresponding corners of the two markers resulting in four lines. Then it measures the angles between each line and horizon (x-axis of image is assumed to be parallel to the true physical horizon) which is shown as θ in Fig. 3. The average of these four angles is considered as the machine-generated angle. Since the AprilTag algorithm has to clearly recognize the boundary of the quadrilateral to finalize the marker detection process, major outliers would not occur. As such, a simple averaging is appropriate to estimate the resultant of the four angles. The same configuration is used for pose estimation of the dipper. The entire architecture of the system is presented in Fig. 4.

FIG. 3: Markers attached to the boom and the recognized angle

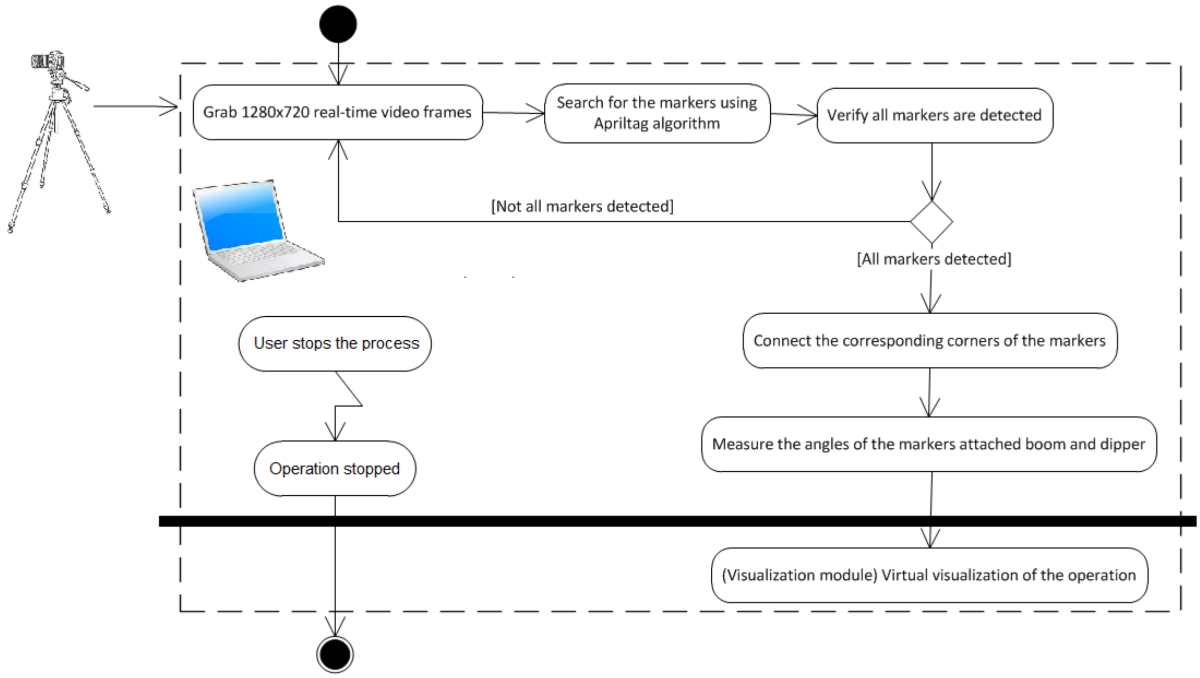

FIG. 4: Architecture of the system

A camera is positioned on the side of the boom where its optical axis is approximately normal to the boom's plane (see Fig. 5). Since user friendliness and quick setup are among the objectives of this system, the camera positioning process was designed to be straightforward while meeting the optical requirements. The users should maintain three main factors during camera setup process: First, the camera should be positioned to look at the middle of arm to cover both boom and dipper. Second, the camera's optical axis should be normal to the arm's plane which could be reached by setting the middle of the arm on the center of the camera's viewfinder. Finally, the camera should be leveled (zero pitch and roll angles) which could be achieved by leveling tubular spirit and bull's eye spirit levels available on regular tripods. This setting does not require camera calibration and meticulous positioning as the system should be able to handle slight off-centered configurations. A sensitivity analysis of the effect of non-normal optical axes and distance is described in the experimental results section. One may argue that an excavator or backhoe's arm rotates to carry and dump the load elsewhere. But the boom pose measurement is only required in a certain yaw angle. The angle would be the same as the angle of the boom's plane and the camera will be faced normal to this plane as shown in Fig. 5. Excavation of a trench is a popular instance for this scenario. If the excavation field changes, the operator simply has to relocate the observing camera accordingly.

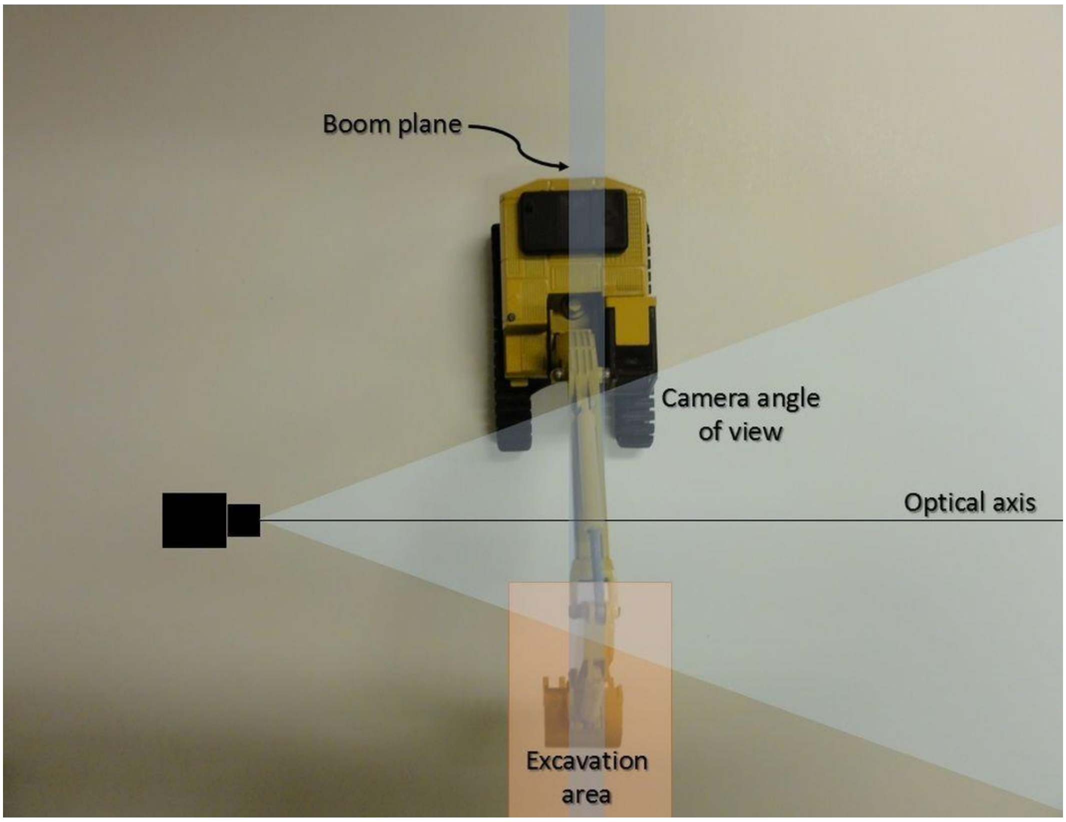

FIG. 5: Camera position relative to excavator and excavation zone

Three main criteria are considered to evaluate this technique: computation time for real-time application, detection rate, and accuracy of the pose measurements. For the first part, a few different popular frame resolutions are processed to evaluate the runtime efficiency on an ordinary personal laptop. For the second and third objectives, machine-generated data are compared to ground-truth data to assess the performance of the framework.

5 EXPERIMENTAL RESULTS

To assess and validate the proposed pose estimation system, we carried out several experiments on actual video sequences. All experiments were processed on a computer with a four-core 2.2 GHz Intel Core i7 CPU, and 8 GB memory. The following subsections present the results for the three experimental criteria including runtime efficiency, detection rate, and accuracy of the pose measurements.

5.1 Runtime efficiency

Two common video sizes, including 640x480 and 1280x720 pixels frames, were processed for which the process times were 0.06 and 0.15 seconds per frame, respectively. This means that the system is able to process about 6.67 frames with 1280x720 pixels resolution in every second (6.67 Hz). This would be within the acceptable range of the industry to monitor excavator's boom as the update rate of existing state of the art excavator gravity-referenced angle sensors are between 2 to 5Hz (Bernold, 2002, Trimble Machine Control http://construction.trimble.com/products/machine-control). Since the AprilTag is a sequential method, the performance of the system mainly depends on the performance of the CPU including clock rate (measured in hertz) and the instructions per clock (IPC). The memory does not have a significant effect as the process used about 0.1 GB of the RAM. In addition, implementation of this algorithm on a graphical processing unit cannot improve runtime efficiency due to sequential nature of the algorithm.

5.2 Detection rate

The first step of this part was to determine the reliable range to detect a marker without any optical or digital zooming. The same resolutions, 640x480 and 1280x720 pixels frames, and two label sizes, including 15.24x15.24 cm (6”x6”) and 20.32x20.32 (8”x8”), were used to determine detection ranges. These marker sizes represent the dimensions of the interior black rectangles which should be surrounded by a white margin for better detection results (see Fig. 1.c). In this experiment, a camera captures videos, each with 230 frames, from various distances to the markers and the reliable range is a distance that the system could detect the marker in more than 98% of the frames (226 frames). The results are presented in Table 1.

TABLE 1. Detection range of AprilTag markers without an optical or digital zoom

Resolution (pixels) / **Marker Size (cm)**	**640 x 480**	**1280 x 720**
15.24x15.24 cm	5.5 m	7.3 m
20.32x20.32 cm	7.5 m	9.5 m

The markers should be small enough to be fitted on the side of most excavators and backhoes. The markers with 20.32x20.32 (8x8 in) apparently provide better range, but they might not be fitted on the small machines with a thin dipper, so 15.24x15.24 cm (6”x6”) was used to proceed with the rest of the tests. The video frame size was set to 1280x720 for the detection rate and accuracy tests, as it has an acceptable detection range (7.3 meter).

5.3 Accuracy of pose estimation

In the next step, two parallel markers were attached on the boom and dipper of a backhoe. The camera was positioned in the six meters range from the backhoe’s boom. This distance is well within the reliable detection range for 15.24x15.24 cm (6”x6”) using 1024x720 pixels (see Table 1). The accurate 3D coordinates of the markers were measured by a total station, which are then used as the ground-truth data. Figure 6 presents the experiment setup.

FIG. 6: Experimental configuration

Three different poses of the backhoe arm were captured for this test, which are shown in Fig. 7. Table 2 shows the results of this experiment.

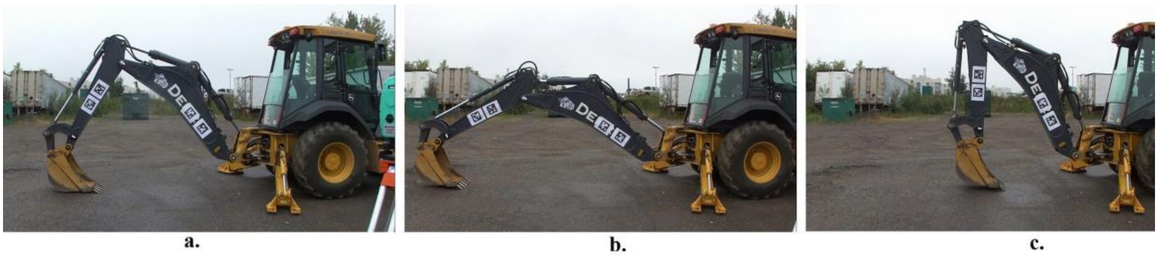

FIG. 7: Boom poses tested in this experiment

TABLE 2. Comparison of ground-truth with machine-generated angle measurements

Pose number	Distance (m)	Arm part	Number of frames	Detection rate (%)	Number of false positives	Average of angle difference (°)	Standard deviation of angle difference (°)	Error propagation of the end point (cm)
Pose 1	6	Boom	79	100%	0	0.161	0.078	2.2
		Dipper	79	100%	0	0.475	0.147	
Pose 2	6	Boom	79	100%	0	0.491	0.153	4.8
		Dipper	79	100%	0	0.479	0.145	
Pose 3	6	Boom	90	100%	0	0.333	0.138	3.5
		Dipper	90	100%	0	0.742	0.146	

The fourth column of this table shows the number of frame in each test, and fifth column presents the detection rate for the labels attached on each arm part which were 100% for all cases. There was no false positive in the processed videos. The seventh column provides the average of the angle difference between the ground-truth measurements (measured by total station) and machine-generated angles. The next column includes standard deviation of the angle difference between machine-generated and ground-truth measurements. The last column presents the error propagation in pose estimation of the endpoint of the arm (excluding bucket) which is calculated using kinematic chain principles assuming a midsize excavator with a 5.6 meter boom and a 2.7 meter dipper. The average error was found to be between 0.161° to 0.742° and the system showed a consistent measurement as the standard deviation of the records were about 0.15°.

5.4 Sensitivity analysis

As mentioned earlier, a main objective of this research project is to develop an easy-to-use system in which positioning of the observer camera should be a quick task with an estimated measurement that the optical axis should be approximately normal to the boom's plane. So the effect of three parameters on the accuracy of angle measurement, including the angle between the camera's optical axis and arm's plane normal, distance between camera and boom's plane, and distance between two tags were investigated.

Altering the angle of the optical axis would result in perspective distortion which increases error in angle measurement. In this experiment, a camera without an optical or digital zooming was positioned at a six meter distance normal to four 15.24x15.24 cm (6"x6") labels which are configured similar to an excavator arm. Then the camera was repositioned with 5 degrees interval on a circle with a radius of six meters to assess its effect on the angle measurement. Figure 8 shows the schematic view of this setting where the angle between the camera's optical axis and arm's plane normal is represented as Θ.

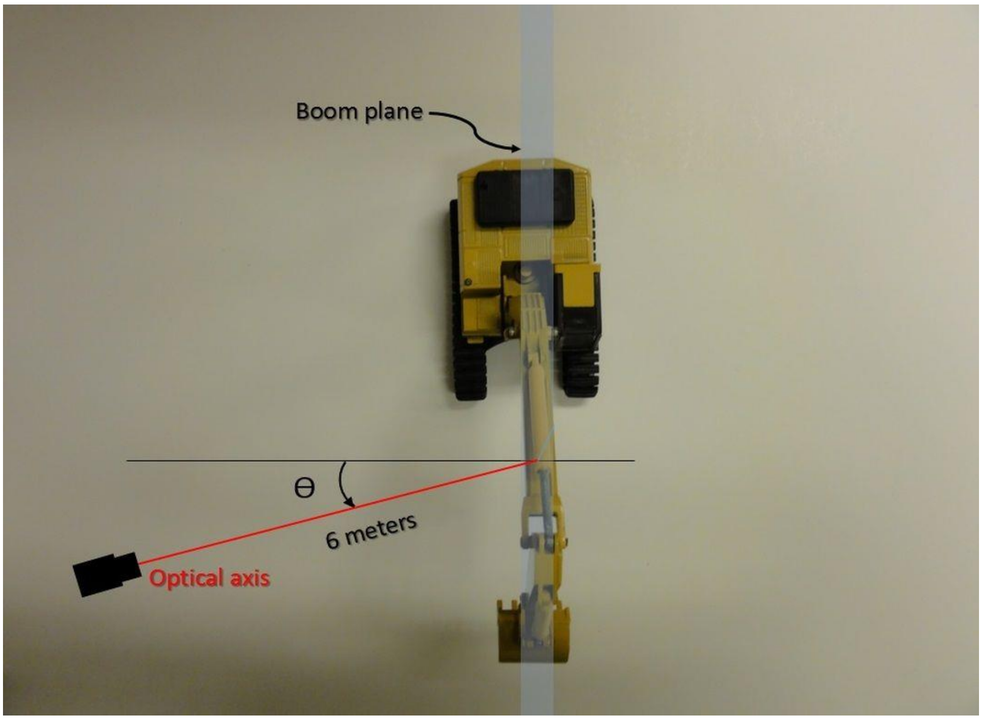

FIG. 8: Configuration of the test for the effect of off-center optical axis

The camera recorded a ten seconds video with five frames per second rate at every position which provided 50 frames at each angle. The four angles for each couple of markers were estimated (similar to previous test) and the average was considered as a machine-generated value which then compared to ground-truth measurements. Table 3 presents the average difference between the averages of machine-generated data and manual angle measurements. The last column shows the error propagation in pose estimation of the endpoint of the arm assuming a midsize excavator with a 5.6 meter boom and a 2.7 meter dipper.

TABLE 3. Sensitivity analysis of optical angle change on machine-generated measurements

Θ angle (°)	average error of the angle measurements (°)	Error propagation of the end point (cm)
0	0.391	3.8
5	0.532	5.2
10	0.935	9.1
15	1.162	11.3
20	1.614	15.7
25	2.520	24.3
30	3.239	31.2

The average errors in the second column show that the increase in the angle between optical axis and arm's plane normal results in larger errors in which the errors up to 5° are not significant. Setting up a camera within the range of -5° to +5° without the need for a precise measurement device is possible in construction sites which could be quickly done with a simple measurement tape which satisfies the easy-to-use objective of this research.

In the second set of tests, a camera without an optical or digital zooming was positioned at half meter intervals starting from 3.5 m to 7.0 m (the maximum distance with reliable detection rate is 7.3 m) normal to four 15.24x15.24 cm (6”x6”) labels which were configured similar to an excavator boom. The camera captured a ten seconds video with five frames per second rate at every position which provided 50 frames at each spot. The average of machine-generated data were compared to ground-truth measurements. The average difference between the averages of machine-generated value and manual angle measurements are provided in Table 4.

TABLE 4. Sensitivity analysis of camera distance on machine-generated measurements

Distance (m)	average error of the angle measurements (°)	Error propagation of the end point (cm)
3.5	0.123	1.2
4.0	0.289	2.8
4.5	0.219	2.1
5.0	0.420	4.1
5.5	0.144	1.4
6.0	0.456	4.4
6.5	0.155	1.5
7.0	0.163	1.6

The results showed no noticeable relationship between the distance (same with focal length) and accuracy of the angle measurements. Increasing the distance for more than 7.3 meter will result in unreliable detection rate and was not investigated.

Lastly, the effect of distance between installed markers was investigated in which the center-to-center distance was increased from 25 cm to 150 cm with 25 cm intervals. Table 5 demonstrates the results. The minimum error was observed in 75 to 100 cm range.

TABLE 5. Sensitivity analysis of distance between two markers

Distance between tags (cm)	average error of the angle measurements (°)	Error propagation of the end point (cm)
25	0.320	3.1
50	0.356	3.5
75	0.122	1.2
100	0.111	1.1
125	0.229	2.2
150	0.244	2.4

5.5 Dynamic measurement

The experiments of the previous section monitored the pose of motionless markers. Although most of the precise excavation or trimming operations are carried out with low speed, the ability of the system to measure the pose of a moving arm has to be investigated as well. Proven high precision measurement devices, such as total stations, only work in static settings and other solutions, including commercial angle sensors, have a certain level of error. As a result, a simple experiment was designed to assess the performance of the system in measurement of the pose of a moving arm.

In this setting, a marked whiteboard (marked every 1°) was installed in background of a rotating arm in which it indicates the angle of the boom in every frame of the video. Due to the large size of an actual excavator arm, the rotating arm was modeled using a smaller arm, but the size of the markers are the same as the ones used in static tests (6”x6”). Figure 9 depicts this setting and also shows the detected markers by the system. A camera was positioned in six meter distance to the arm without any optical or digital zooming. Since the size and distance of the markers to the camera are the same, the size and type of the rotating arm would not have a significant effect compared to an actual machine.

FIG. 9: Experiment setting for dynamic measurement

The camera recorded the movement of the labeled arm with the rate of five frames per second. The frames resolution was set to 1280x720 pixels and as the system could process about 6.67 frames per second, so it could maintain the real-time video stream. Four videos where captured in which the arm was rotated with four angular speeds: 0.134 rad/sec, 0.23 rad/sec, 0.417 rad/sec, and 0.535 rad/sec. These angular speeds are within the range of regular excavator operation. More specifically this system is intended to monitor precise excavation or trimming operations which are carried out carefully with slow movement to achieve the design profiles. The results of the dynamic measurement are provided in Table 6.

TABLE 6. Results of the dynamic tests

Angular Speed (rad/sec)	Distance (m)	Number of frames	Detection rate (%)	Number of false positives	Average of angle difference (°)	Standard deviation of angle difference (°)	Maximum angle difference (°)
0.134	6	139	100%	0	0.908	0.242	1.339
0.23	6	104	100%	0	0.856	0.321	1.375
0.417	6	49	100%	0	0.907	0.417	1.435
0.535	6	44	98.86%	0	0.871	0.353	1.435

Both averages of error and standard deviations are larger in dynamic tests than in static experiment. These outcomes, however, are measured using visual scale and are not as accurate as the total station. The visual scale was marked every 1° in measurement and considering the translation of actual dimensions to the pixels in 6 meter distance, 0.5° error is possible in every measurement. So, the results of these tests are not as reliable as the static ones and as mentioned before, using accelerometer or gravity-referenced angle sensors cannot provide much better ground-truth data.

The algorithm is able to detect and estimate the pose of markers robustly, and its performance does not relate to the speed of the markers. In fact, the performance depends on the quality of the camera to capture motion and provide clear video sequences for more accurate results.

6 DISCUSSION

This measurement system has a kinematic equivalency problem which is shown in Fig. 10. An angular offset between the actual rotation axis (shown by solid line in Fig. 10) and measured axis by the system (shown by dashed line in Fig. 10) should be accounted for. This issue is more applicable to the boom due to its special curves, and the angular offset is not necessary for most of the dippers as the labels could be installed parallel to the rotation axis (same as the case in Fig. 10). In the scenario of the boom which is illustrated in Fig. 10, the angular offset should be subtracted from the measured orientation. The angular offset was 10.56° for the test backhoe (in Fig. 10).

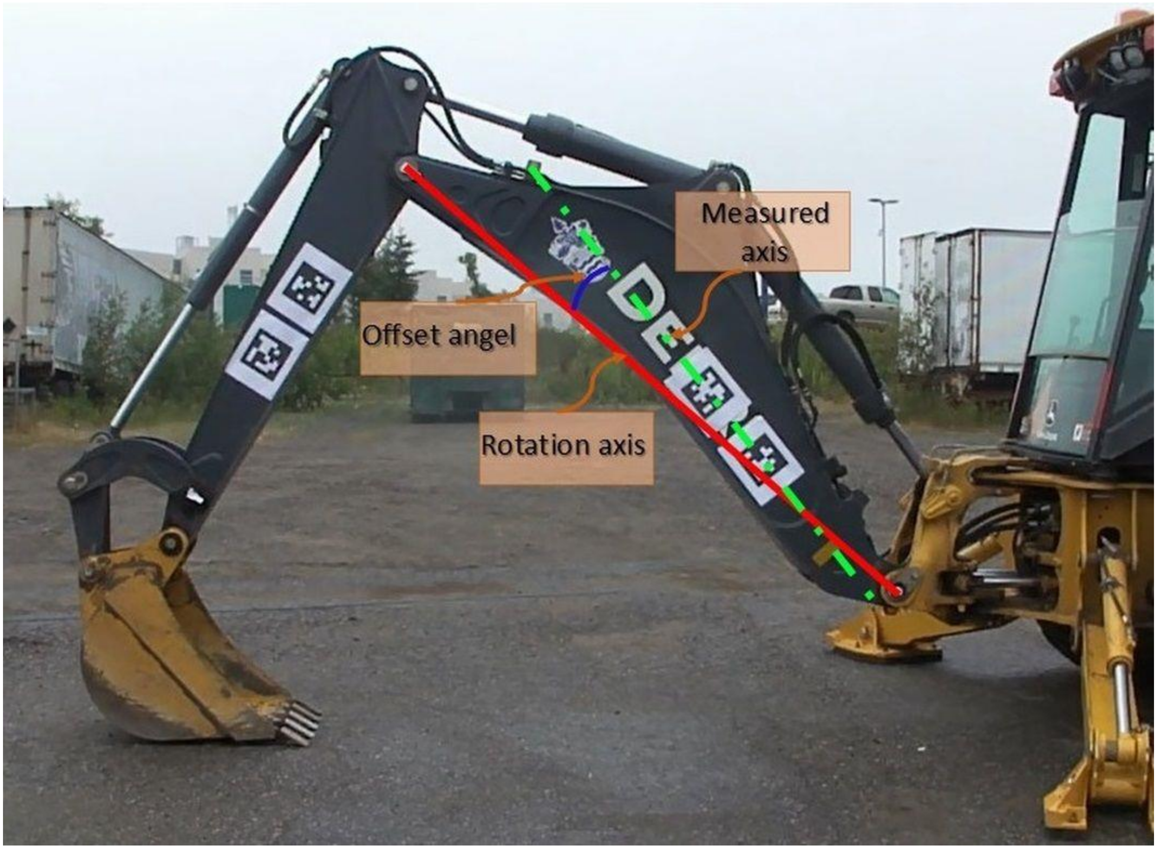

FIG. 10: Angle difference between measured and rotation axes

The average errors were between 0.161° to 0.742° in static tests. Considering a regular midsize excavator with a 5.6 meter boom (rotation axis) and a 2.7 meter dipper (rotation axis), the angular errors could be translated to the measurement error of the end of the dipper. The kinematic chain principle was used to calculate the error of the end of the chain while the rotation angles for the boom and the dipper varied from 0° to 90° and 0° to -90°, respectively. These angles are depicted in Fig. 11. The extreme scenario was considered in the calculations where both the measurements of the both boom and dipper had 0.742° angular error. The outcomes of these analysis showed that the worst combination results in 7.2 centimeter error in depth estimation of the end of the kinematic chain which is the pin connecting the dipper to the bucket. In a related research project which was developed using electronic joint encoders and laser receivers, 5.1 centimeter accuracy was achieved in the field experiments (Bernold, 2002).

Given the angle measurements (the vision-based method, as tested, is not able to measure the angle of bucket), 3D model CAD data of the exactor, and GIS CAD data of the excavation profile, a hybrid virtual reality system can emulate the movement of the boom in real time and provide the distance between the end effector and the excavation level in a graphical display in the cabin for the operator (see Fig. 11). The implementation of this hybrid virtual system has been explained in detail in prior work (Talmaki and Kamat 2014).

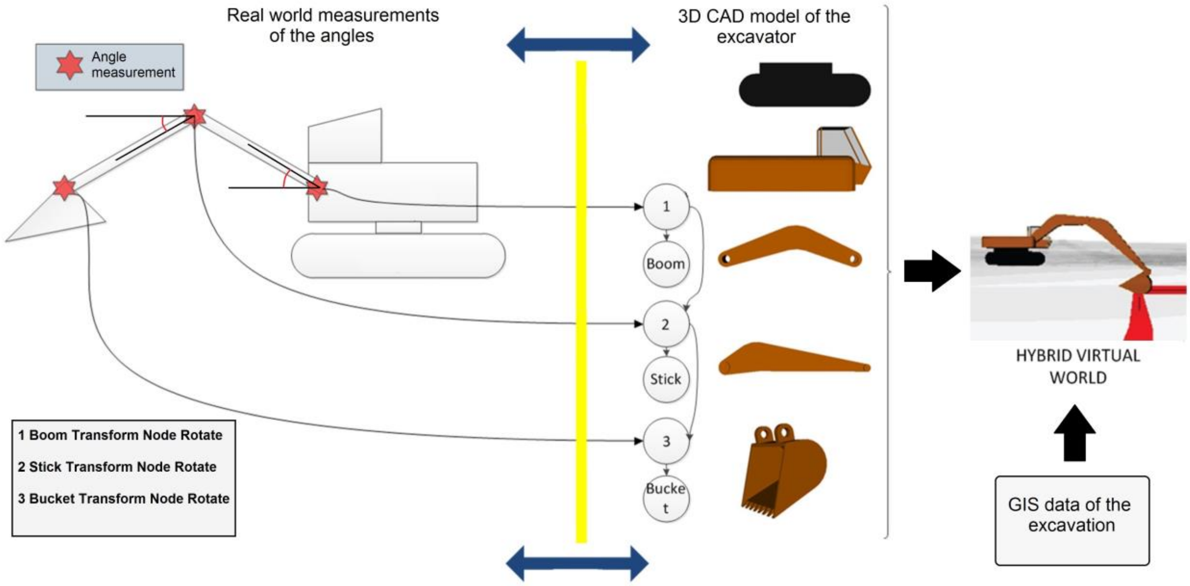

FIG. 11: Real-world angle measurements + 3D Cad model of the machine + GIS data of excavation updating hybrid virtual world

The experimental results showed the promising performance of the system in the measurement of the boom pose that could fulfill the expectations of many earthwork operations. This system requires a regular digital camera, a tripod, few markers, and a processing system, such as a laptop or tablet. The total cost of these components could vary from $1700 to $2000, based on the type of the camera and processor. In addition, acceleration (such as sudden stop) is not an issue in this approach whereas the systems using tilt sensors require filtering the sensor signals. This measurement technique, however, has one main shortcoming which is the inability to estimate the pose of an end-effector such as a bucket. This issue might also happen for the dipper in the cases where the excavator has to dig deep trenches or dig under water in which cases the marker disappears from the camera's field of view. The competitive advantages and shortcomings of this system, and external affecting factors are summarized in Fig. 12 using SWOT analysis.

Strengths:	Weaknesses:
• User-friendly • Easy to set up • Low-cost	• Unable to track the bucket • Unable to track 3D pose • Requires clear line-of-sight • Requires camera plane to be parallel to marker plane • Requires camera to be leveled
Opportunities:	**Threats:**
• Tendency to improve excavation productivity • Large number of small companies or occasional users • Large market for low-cost solutions in developing countries	• Existing sensor-based solutions for tracking 3D pose of all components • Slow rate of technology adoption in construction

FIG. 12: SWOT analysis of the system

The developed system could provide more pose data than 1D digging control systems, such as laser-based depth estimation systems (http://www.leica-geosystems.com/en/Machine-Control_4677.htm, http://www.topconpositioning.com/products/machine-control) which only estimate the elevation of the dipper. But it has inferior performance compared to 2D control systems as it cannot provide orientation of the bucket and also might lose tracking markers due to obstruction. There is an ongoing research to develop a low-cost solution to measure the pose of the bucket for this system.

7 CONCLUSION

The competitive environment and low profit margins of the earthwork market have created a large demand for more productive solutions. Machine control technologies are among the growing products to address this issue which have been beneficial in large projects. High prices of these products, however, have limited their application in small and occasional operations. This paper presents a new framework which uses a state-of-the-art marker-based recognition algorithm to measure the pose of an excavator's arm. This system requires a regular processing platform, a camera, and four low-cost visual markers to monitor the 2D pose of the boom and dipper of an excavator or a backhoe.

In this framework, the AprilTag algorithm was used to detect parallel markers attached to a backhoe's boom and dipper. Then the average of the angles between marker sides and horizon were calculated as the angle of the arm element. The average errors in experiments on an actual backhoe were between 0.161° to 0.742°. The most severe case, where both the boom and dipper angles have 0.742° error, would result in 7.2 centimeter error in depth estimation of a midsize excavator's arm. This system, however, has some shortcomings that include the inability to measure the pose of the bucket, and the dipper in deep trenches, as well as the yaw and roll angle of the arm plane. These challenges are currently being addressed by the authors via full 6 degree of freedom marker-based pose (position and orientation) estimation to enable accurate depth tracking of the bucket.

REFERENCES

Bernold L.E. (2002). Spatial integration in construction, *Journal of Construction Engineering and Management*, Vol. 128, No.5, 400–408.

Bernold L.E. (2007). Control schemes for tele-robotic pipe installation, *Automation in Construction*, Vol. 16, No.4, 518-524.

Chi S. and Caldas C. H. (2011). Automated object identification using optical video cameras on construction sites, *Computer-Aided Civil and Infrastructure Engineering*, Vol. 26, No. 5, 368–380.

Chiang M.H. and Huang C.C. (2004). Experimental implementation of complex path tracking control for large robotic hydraulic excavators, *The International Journal of Advanced Manufacturing Technology*, Vol. 23, No. 1-2, 126-132.

Dollar P., Wojek C., Schiele B. and Perona P. (2012). Pedestrian detection: an evaluation of the state of the art, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 34, No. 4, 743 - 761.

Felzenszwalb P.F., Girshick R.B., McAllester D. and Ramanan D. (2010). Object detection with discriminatively trained part based models, *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol. 32, No. 9, 1627 – 1645.

Feng C. and Kamat V.R. (2013). Plane registration leveraged by global constraints for context-aware AEC applications, *Computer-Aided Civil and Infrastructure Engineering*, Vol. 28, 325–343.

Fiala M. (2005). ARTag, a fiducial marker system using digital techniques, *in Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) San Diego, CA, USA*, Vol. 2, 590–596.

Gong J. and Caldas C.H. (2011). An object recognition, tracking, and contextual reasoning-based video interpretation method for rapid productivity analysis of construction operations, *Automation in Construction*, Vol. 20, No. 8, 1211–1226.

Han S., Lee S., Hong T. and Chang H. (2006). Simulation analysis of productivity variation by global positioning system implementation in earthmoving operations, *Canadian Journal of Civil Engineering*, Vol. 33, No. 9, 1105-1114.

Huang X. and Bernold L. (1997). CAD-integrated excavation and pipe laying. *Journal of Construction Engineering and Management*, Vol. 123, No.3, 318–323.

Kato H. and Billinghurst M. (1999). Marker tracking and hmd calibration for a video-based augmented reality conferencing system, *in Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality*, 85-94.

Lee J., Lorenc S. and Bernold L. (2003). Comparative performance evaluation of tele-operated pipe laying, *Journal of Construction Engineering and Management*, Vol. 129, No.1, 32–40.

Meguro J., Murata T., Takiguchi J., Amano Y. and Hashizume T. (2009). GPS multipath mitigation for urban area using omnidirectional infrared camera, *IEEE Transactions on Intelligent Transportation Systems*, Vol. 10, No. 1, 22-30.

Memarzadeh M., Golparvar-Fard M. and Niebles J.C. (2013). Automated 2D detection of construction equipment and workers from site video streams using histograms of oriented gradients and colors, *Automation in Construction*, Vol. 32, 24-37.

Navon R., Goldschmidt E. and Shpatnisky Y. (2004). A concept proving prototype of automated earthmoving control, *Automation in Construction*, Vol. 13, No. 2, 225– 239.

Navon R. and Shpatnisky Y. (2005). Field experiments in automated monitoring of road construction, *Journal of Construction Engineering and Management*, Vol. 131, No. 4, 487– 493.

Olson E. (2011). AprilTag: a robust and flexible visual fiducial system, *in Proceedings of IEEE International Conference on Robotics and Automation (ICRA'11) Shanghai, China*, 3400–3407.

Rezazadeh Azar E. and McCabe B. (2012a). Part based model and spatial-temporal reasoning to recognize hydraulic excavators in construction images and videos, *Automation in construction*, Vol. 24, 194-202.

Rezazadeh Azar E. and McCabe B. (2012b) Automated visual recognition of dump trucks in construction videos, *Journal of computing in civil engineering*, Vol. 26, No. 6, 769-781.

Rezazadeh Azar E., Dickinson S. and McCabe B. (2013). Server-customer interaction tracker: computer vision–based system to estimate dirt-loading cycles, *Journal of Construction Engineering and Management*, Vol. 139, No. 7, 785–794.

Stentz A., Bares J., Singh S. and Rowe P. (1999). A robotic excavator for autonomous truck loading, *Autonomous Robots*, Vol. 7, No. 2, 175-186.

Talmaki S. and Kamat V.R. (2014). Real-time hybrid virtuality for prevention of excavation related utility strikes, *Journal of Computing in Civil Engineering*, Vol. 28, No. 3, 04014001.

Teizer J., Venugopal M. and Walia A. (2008). Ultrawideband for automated real time three-dimensional location sensing for workforce, equipment, and material positioning and tracking, *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2081, 56-64.

Yamamoto H., Moteki M., Shao H., Ootuki T., Kanazawa H. and Tanaka Y. (2009). Basic technology toward autonomous hydraulic excavator, in *26th International Symposium on Automation and Robotics in Construction (ISARC 2009) in Austin, USA*, 288-295.

Yang Y. and Ramanan D. (2011). Articulated pose estimation with flexible mixtures-of-parts, in *Computer Vision and Pattern Recognition (CVPR) 2011 Conference in Providence, RI, USA*, 1385 - 1392.