

NCHRP IDEA Program

Exploratory Analysis of Augmented Reality Visualization for Right-of-Way Excavation Safety

Final Report for
NCHRP IDEA Project 167

Prepared by:
Vineet R. Kamat, Ph.D.
University of Michigan

July 2016

Innovations Deserving Exploratory Analysis (IDEA) Programs Managed by the Transportation Research Board

This IDEA project was funded by the NCHRP IDEA Program.

The TRB currently manages the following three IDEA programs:

- The NCHRP IDEA Program, which focuses on advances in the design, construction, and maintenance of highway systems, is funded by American Association of State Highway and Transportation Officials (AASHTO) as part of the National Cooperative Highway Research Program (NCHRP).
- The Safety IDEA Program currently focuses on innovative approaches for improving railroad safety or performance. The program is currently funded by the Federal Railroad Administration (FRA). The program was previously jointly funded by the Federal Motor Carrier Safety Administration (FMCSA) and the FRA.
- The Transit IDEA Program, which supports development and testing of innovative concepts and methods for advancing transit practice, is funded by the Federal Transit Administration (FTA) as part of the Transit Cooperative Research Program (TCRP).

Management of the three IDEA programs is coordinated to promote the development and testing of innovative concepts, methods, and technologies.

For information on the IDEA programs, check the IDEA website (www.trb.org/idea). For questions, contact the IDEA programs office by telephone at (202) 334-3310.

IDEA Programs
Transportation Research Board
500 Fifth Street, NW
Washington, DC 20001

The project that is the subject of this contractor-authored report was a part of the Innovations Deserving Exploratory Analysis (IDEA) Programs, which are managed by the Transportation Research Board (TRB) with the approval of the National Academies of Sciences, Engineering, and Medicine. The members of the oversight committee that monitored the project and reviewed the report were chosen for their special competencies and with regard for appropriate balance. The views expressed in this report are those of the contractor who conducted the investigation documented in this report and do not necessarily reflect those of the Transportation Research Board; the National Academies of Sciences, Engineering, and Medicine; or the sponsors of the IDEA Programs.

The Transportation Research Board; the National Academies of Sciences, Engineering, and Medicine; and the organizations that sponsor the IDEA Programs do not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the object of the investigation.

**Exploratory Analysis of Augmented Reality Visualization for
Right-of-Way Excavation Safety**

**IDEA Program Final Report
NCHRP IDEA Project 167**

Prepared for the IDEA Program
Transportation Research Board
The National Academies

Submitted by

Vineet R. Kamat, Ph.D. (PI)
Department of Civil and Environmental Engineering
University of Michigan
2350 Hayward Street, Suite 2340 G.G. Brown Building
Ann Arbor, Michigan 48109
vkamat@umich.edu
(734) 764-4325

July 2016

ACKNOWLEDGMENTS

This project was funded by the Transportation Research Board of the National Academies under the TRB-IDEA Program (NCHRP-167). The research team extends its sincere gratitude to the TRB Project Director Dr. Inam Jawed for his continuous guidance and encouragement throughout the project. The research team is also grateful to the members of the IDEA advisory panel, David Kuehn from the Department of Transportation and Valerie Shuman from the Shuman Consulting Group, for their feedback and patient guidance throughout the project. Thanks are also extended to members of the regional advisory board, James DeKimpe from the DTE Energy Company and Eric Urbain from MISS DIG System, for their cooperation and support during this project. The research team also thanks Walbridge Construction Company, Eagle Excavation Company, and the University of Michigan Architecture, Engineering, and Construction (AEC) division for their support in providing access to construction equipment and job sites for experimentation and demonstration.

NCHRP IDEA PROGRAM COMMITTEE

CHAIR

DUANE BRAUTIGAM

Consultant

MEMBERS

CAMILLE CRICHTON-SUMNERS

New Jersey DOT

AGELIKI ELEFTERIADOU

University of Florida

ANNE ELLIS

Arizona DOT

ALLISON HARDT

Maryland State Highway Administration

JOE HORTON

California DOT

MAGDY MIKHAIL

Texas DOT

TOMMY NANTUNG

Indiana DOT

MARTIN PIETRUCHA

Pennsylvania State University

VALERIE SHUMAN

Shuman Consulting Group LLC

L.DAVID SUITS

North American Geosynthetics Society

JOYCE TAYLOR

Maine DOT

IDEA PROGRAMS STAFF

STEPHEN R. GODWIN

Director for Studies and Special Programs

JON M. WILLIAMS

Program Director, IDEA and Synthesis Studies

INAM JAWED

Senior Program Officer

DEMISHA WILLIAMS

Senior Program Assistant

EXPERT REVIEW PANEL

DAVID KUEHN, *FHWA*

VALERIE SHUMAN, *Shuman Consulting Group*

JAMES DEKIMPE, *DTE Energy Company*

ERIC URBAIN, *MISS DIG System*

FHWA LIAISON

DAVID KUEHN

Federal Highway Administration

TRB LIAISON

RICHARD CUNARD

Transportation Research Board

COOPERATIVE RESEARCH PROGRAM STAFF

STEPHEN PARKER

Senior Program Officer

TABLE OF CONTENTS

1	EXECUTIVE SUMMARY	1
2	IDEA PRODUCT	2
3	CONCEPT AND INNOVATION	3
3.1	Lack of Persistent Visual Guidance	3
3.2	Inability to Gauge Proximity of Digging Implement to Buried Assets	4
3.3	Augmented Reality Visualization of Geospatial Utility Data	5
4	INVESTIGATION	5
4.1	Geometric Registration of Geospatial Utility Data	5
4.1.1	Resolving the Latency Problem in Orientation Sensors	9
4.1.2	Adaptive Latency Compensation Algorithm	12
4.1.3	Field Demonstration of Augmented Reality Registration Algorithms	12
4.2	Real-Time Pose Estimation of Articulated Excavator	16
4.2.1	Camera Marker Network	18
4.2.2	Experimental Prototypes	20
4.2.3	System Calibration	21
4.2.4	Uncertainty Analysis	233
4.2.5	Uncertainty Propagation	23
4.2.6	Uncertainty and Configuration	24
4.2.7	Feasibility Experiments	26
4.2.8	Field Demonstration of Articulated Machine Pose Estimation System	300
5	PLANS FOR IMPLEMENTATION	333
6	CONCLUSIONS	344
	REFERENCES	366

1 EXECUTIVE SUMMARY

An excavator unintentionally hits a buried utility every 60 seconds in the United States, causing billions of dollars in damage each year. Most of these accidents occur along public rights-of-way (ROWs). Typically, these accidents occur either because excavator operators do not know where utilities are buried (inaccurate or missing utility location data) or because they cannot perceive where the utilities are relative to the digging excavator (inaccurate spatial perception). This IDEA project addressed these problems by exploring new methods to communicate utility location data to equipment operators during excavation. The research created and evaluated two key capabilities: (1) persistent visualization of assets buried in an excavator's vicinity using a georeferenced augmented reality (AR) approach; and (2) real-time monitoring of an excavator's proximity to underground utilities using a graphical emulation approach. These unique capabilities can allow an excavator operator to be visually aware of buried assets in a machine's vicinity and provide a real-time quantitative measure of a machine's distance to nearby obstructions, offering the potential to significantly reduce the occurrence of buried utility strikes in ROW excavation.

In this report, the results of the project are described in two stages. First, a graphics algorithm was designed to place virtual entities in an augmented scene given an excavator's geographical position and orientation. The algorithm's accuracy was validated in static scenarios of inspection of underground utilities. Stabilization approaches were designed to reduce the jittering response of orientation tracking data, caused by the electronic compass that is vulnerable to vibrations and magnetic interference. The convention of marking underground utilities was studied and an Augmented Reality visualization system compatible with diverse geospatial data formats was designed. Visualization methods that obstruct the ground view, while rendering utilities underneath, were also implemented. A "first-person" operator-view Augmented Reality visualization system was subsequently designed and evaluated using known subsurface utility geospatial data records.

The second stage of the project focused on practical implementation of these algorithms on excavators in the field. Prior technologies in place to determine the pose of an articulated excavator were investigated, and the limitations in current methods were analyzed. The pose of an articulated machine includes the position and orientation of not only the machine base (e.g., tracks or wheels), but also each of its major articulated components (e.g., stick and bucket). A computer vision-based solution using a network of cameras and markers was designed to enable such a capability for articulated machines. A planar marker is magnetically mounted on the stick (dipper) of an excavator. Another marker is fixed on the job site whose 3D pose is pre-surveyed in a project coordinate frame. Then a cluster of at least two cameras respectively observing and tracking the two markers simultaneously forms a camera-marker network and transfers the excavator's pose into the desired project frame, based on a pre-calibration of the relative poses between each pair of cameras. Through extensive sets of uncertainty analyses and field experiments, this approach was shown to be able to achieve centimeter level tracking accuracy within 10 meters with only two ordinary cameras and a few markers. A working prototype was tested on several active construction sites with positive feedback from excavator operators confirming the solution's effectiveness.

In addition to promising research results, the project identified several avenues for future work aimed at facilitating large-scale deployment of the technology in the field. Improving the accuracy and the footprint of the technology components were identified as key areas of focus for future development efforts. The system accuracy has a quasi-linear relationship to the marker size. While the current marker size is practical for daily transportation and installation, any desired shrinkage in marker size can be compensated for by longer focal length lenses or optimized stereo marker designs. A network of multiple base markers, whose relative transformation to each other are pre-calibrated with bundle adjustment, may also be necessary in projects with large horizontal or linear expanses. The footprint of such networks can possibly cover entire sites so that the base camera can pick one base marker within its working radius from any location.

2 IDEA PRODUCT

The IDEA product developed in this project is capable of using AR visualization and emulated graphical monitoring to estimate an excavator's proximity to "invisible" buried utilities in real-time while the machine is digging. The technology can allow an excavator operator to be visually aware of buried underground utilities and other assets in a machine's vicinity, and can offer quantitative feedback of the machine's distance to obstructions in the vicinity, thereby preventing accidental utility strikes and improving productivity (Figure 1).

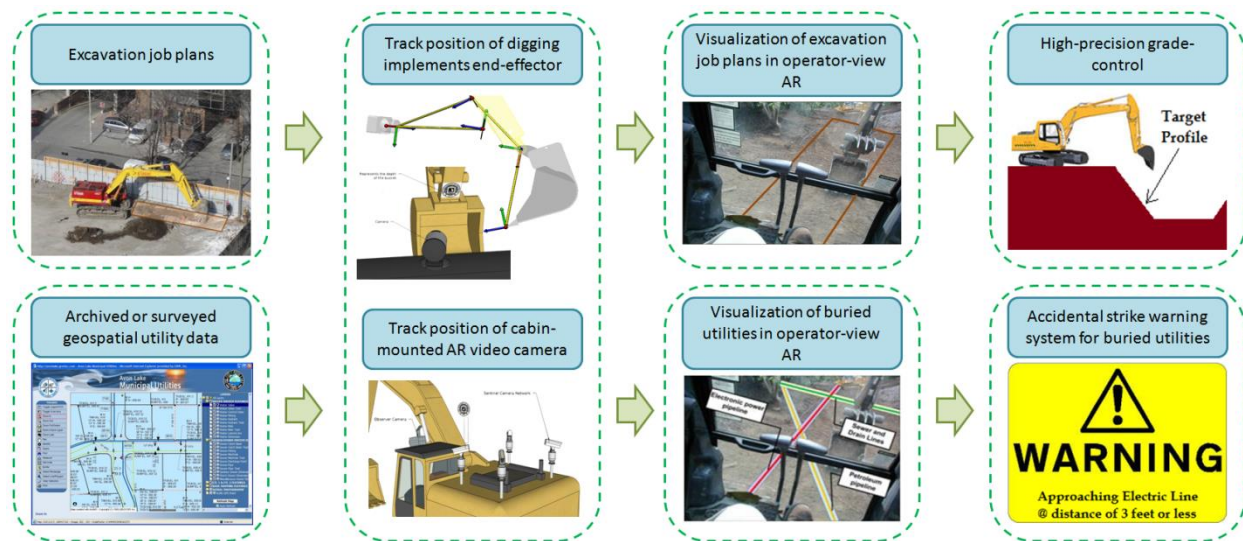


Figure 1. Augmented reality visualization of geospatial utility data:
Precise grade-control (above) and utility avoidance (below).

The technical success of this product was contingent on the development of two main features: (1) Tracking the position of the digging excavator's implement (bucket) for high-precision grade-control; and (2) Tracking the

position of the cabin-mounted AR video-feed camera for visualizing buried utilities in operator-view AR and proximity monitoring for utility strike warnings. The core disruptive innovation in the solution is the computer-vision-based motion-tracking platform technology that is used to track a working excavator's motion in a local coordinate system in order to register buried utility locations for persistent AR visualization. The developed motion-tracking system uses visual markers, calibrated camera networks, and computer-vision algorithms to accurately track the position of key excavator components relative to a local coordinate system defined by a benchmark visual marker.

Accidents involving excavator hits to utilities are a long-standing and significant societal problem that leads to an unacceptable number of fatalities, injuries, property damage, and other costs each year (Sterling et al. 2009; Spurgin et al. 2009). Inadvertent utility strikes disrupt life and commerce, and pose physical danger to workers, bystanders, and the general public. The explored innovations thus have the potential to transform excavator operation from a skill-based to a knowledge-based process so that future accidents are prevented.

3 CONCEPT AND INNOVATION

An excavator unintentionally hits a buried utility every 60 seconds in the United States, causing hundreds of fatalities, thousands of injuries, and billions of dollars in property damage and environmental pollution each year (Spurgin et al. 2009). Most of these accidents occur along public rights-of-way (ROWs) such as roadways and sidewalks. Typically, these accidents occur either because excavator operators do not know where utilities are buried or because they cannot perceive where the utilities are located relative to the digging excavator (Anspach 2011). Inaccurate, incomplete, or missing utility location information is often cited as a cause of incidents involving excavators striking buried utilities (Sterling et al. 2009; Patel et al. 2010; Anspach 2011). However, despite advances in geophysical locating technologies, and independent of the care with which known utility locations are marked on the ground, there are some fundamental challenges that make it very difficult for excavator operators to be spatially aware of their surroundings while digging. The state of practice has two critical limitations when considered from an excavator operator's perspective:

- Lack of persistent visual guidance
- Inability to gauge proximity of excavator to buried assets.

3.1 LACK OF PERSISTENT VISUAL GUIDANCE

While the practice of marking utility locations on the ground helps in initial excavation planning, there is a basic practical limitation: The surface markings (paint, stakes, flags, etc.) are the first things to be destroyed or dislocated when excavation begins and the top soil or surface is scraped. This makes it challenging for an excavator operator to maintain spatial orientation (Figure 2), who must then rely on memory and judgment to recollect expected utility locations as excavation proceeds. Seemingly innocuous events such as returning from lunch or stopping to help another crew can prove detrimental because the operator must now recall the marked locations before continuing to

dig. Thus, excavator operators and field supervisors have no persistent visual clues that can help them be spatially aware of the underground space surrounding an excavator's digging implement.



Figure 2. Excavator operators' perspective.

3.2 INABILITY TO GAUGE PROXIMITY OF DIGGING IMPLEMENT TO BURIED ASSETS

Another significant limitation is that an operator has no practical means of knowing the distance of an excavator's digging implement (e.g., bucket) to the nearest buried obstructions until they are exposed. Excavation guidelines in most states require buried utilities to be hand-exposed prior to using power equipment (MDS 2009). Failure to follow the hand-exposure guidelines, which happens often out of ignorance or as a conscious decision, means that the first estimate of proximity an operator receives is when the digging implement actually “touches” a buried utility. It is easy to understand why this first “touch” can often actually be a “strike.”

Field locators and markers typically mark only the horizontal location of utilities on the ground, and often do not include any depth or other attribute information that may possibly help operators better perceive their location in three dimensions (MDS 2009). A possible justification for this practice is that locators believe that there are “standard” depths where each utility type is typically buried. Second, even if depth information is marked on the ground along with a utility's horizontal location, the markings are destroyed early in the excavation process, placing

the burden of remembering a utility's expected depth and orientation on the operator. Thus, without any visual cues or quantitative feedback, operators find it very challenging to gauge the evolving distance between a digging machine and any unexposed utilities buried in the vicinity.

3.3 AUGMENTED REALITY VISUALIZATION OF GEOSPATIAL UTILITY DATA

This IDEA project explored a new technical approach that uses AR visualization of subsurface utility geospatial data to enable knowledge-based excavator control. In AR, computer-generated images are superimposed over a user's view of the real world. By presenting contextual information in AR, it is possible to enhance or augment the user's view beyond the normal experience, thus being of possible assistance in the performance of several engineering tasks. The key challenge for AR visualization is registration of superimposed graphics onto the real background, such that the two provide an illusion of co-existence (Azuma et al. 2001). In this case, it means that as an excavator digs, the superimposed geospatial data graphics must stay "fixed" to their intended ground locations to continuously help orient the operator. The specific innovations necessary to implement this concept include algorithms to accurately track the arbitrary motion of a working excavator, and register superimposed geospatial data graphics into the excavator operator's view for improved spatial awareness of the machine's surroundings (Figure 3).

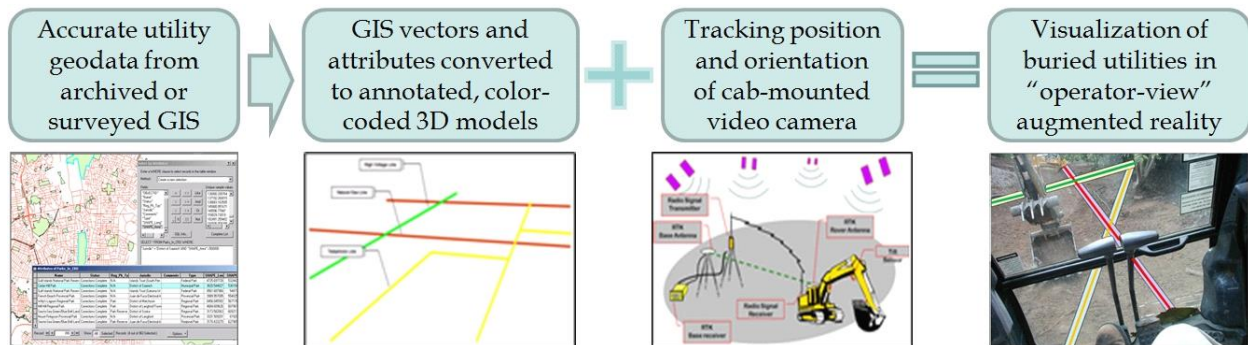


Figure 3. Technical approach for visualizing geospatial data in operator-view augmented reality.

4 INVESTIGATION

The overall design and development in this project was pursued in two inter-related stages:

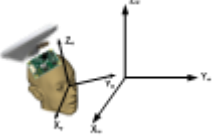
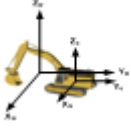


1. Graphical registration techniques to visualize buried utility geospatial data in operator-perspective AR.
2. Real-time methods to track the pose of an articulated excavator relative to known utility locations.

4.1 GEOMETRIC REGISTRATION OF GEOSPATIAL UTILITY DATA

This design and development stage addressed the computational details of geometric registration of geospatial utility data for 3D visualization. The registration process of augmented reality is very similar to the computer graphics transformation process: (1) positioning the viewing volume of a user's eyes in the world coordinate system; (2) positioning objects in the world coordinate system; (3) determining the shape of the viewing volume; and (4)

converting objects from the world coordinate system to the eye coordinate system. However, unlike computer graphics where parameters needed for step 1 through 3 are coded or manipulated by the user, augmented reality rigidly enforces these steps according to the 6 degrees of freedom measured by tracking devices and the lens parameters of the used camera. Table 1 lists the developed registration process, parameters, and measuring devices.

Table 1. The Four Steps of the Developed Registration Process

Step	Task	Illustration	Parameters and Device
Viewing	Position the viewing volume of a user's eyes in the world		Attitude of the camera (electronic compass)
Modeling	Position the objects in the world		Location of the world origin (RTK GPS)
Creating Viewing Frustum	Decide the shape of the viewing volume		Lens and aspect ratio of camera (camera)
Projection	Project the objects onto the image plane		Perspective Projection Matrix

Step 1—Viewing: The origin of the world coordinate system coincides with that of the eye coordinate system, which is the user's geographical location (Figure 4). The world coordinate system uses a right-handed system with the Y-axis pointing in the direction of the true north, the X-axis pointing to the east, and the Z-axis pointing upward. The eye coordinate system complies with the OpenSceneGraph (OSG) default coordinate system, using a right-handed system with the Z-axis as the up vector, and the Y-axis departing from the eye.

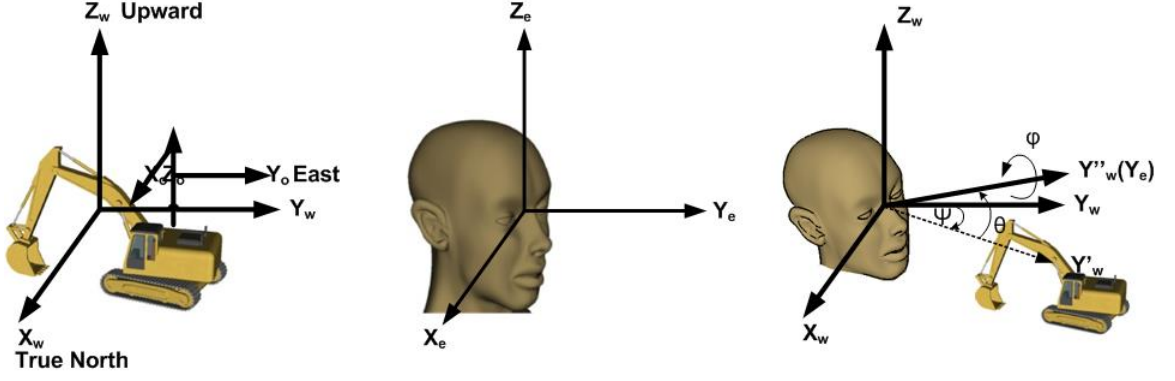


Figure 4. Definition of the world coordinate system.

Yaw, pitch, and roll—all measured by an orientation sensor—are used to describe the relative orientation between the world and eye coordinate systems (Figure 5). There are six possibilities of rotating sequences (i.e., xyz, xzy, zxy, zyx, yzx, and yxz), and zxy is picked to construct the transformation matrix between the two coordinate systems.

Suppose the eye and world coordinate systems coincide at the beginning; the user's head rotates around the Z-axis by yaw angle $\Psi \in (-180, 180]$ to get the new axes, X' and Y' . Since the rotation is clockwise under the right-handed system, the rotation matrix is $R_z(-\Psi)$. Secondly, the head rotates around the X' -axis by pitch angle $\theta \in [-90, +90]$ to get the new axes, Y'' and Z' , with counter-clockwise rotation of $R_{x'}(\theta)$. Finally, the head rotates around the Y'' -axis by roll angle $\phi \in (-180, 180]$ with a counter-clockwise rotation of $R_{y''}(\phi)$ to reach the final attitude.

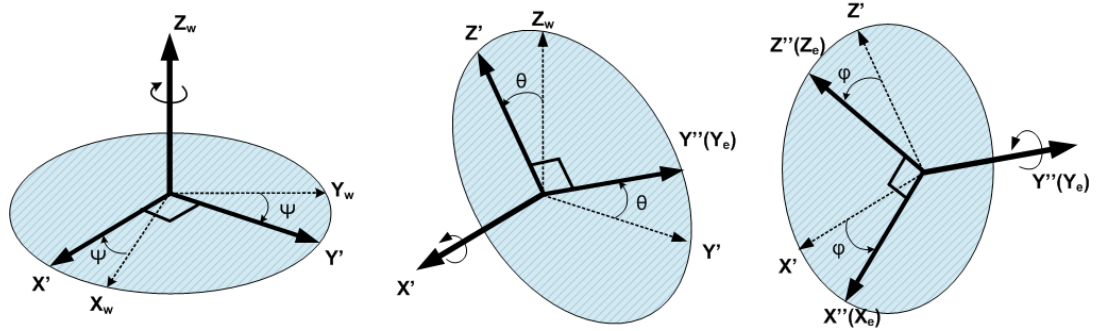


Figure 5. Relative orientation between world and eye coordinates as yaw, pitch, and roll.

Converting the virtual object from the world coordinate to the eye coordinate is an inverse process of rotating from the world coordinate system to the eye coordinate system, therefore the rotating matrix is written as: $R_z(\Psi) R_{x'}(-\theta) R_{y''}(-\phi)$ or $R_z(\text{yaw}) R_{x'}(-\text{pitch}) R_{y''}(-\text{roll})$ (Equation 2). Since OSG provides quaternion, a simple and robust way to express rotation, the rotation matrix is further constructed as quaternion by specifying the rotation axis and angles. The procedure is explained as follows, and its associated equations are listed in sequence from Equation 3 to 5:

rotating around the Y'' -axis by $-\varphi$ degree, then rotating around the X' -axis by $-\theta$ degree, and finally rotating around the Z -axis by Ψ degree.

$$\begin{bmatrix} Xe \\ Ye \\ Ze \end{bmatrix} = \begin{bmatrix} \cos\Psi & \sin(-\Psi) & 0 \\ \sin\Psi & \cos\Psi & 0 \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & \sin(-\theta) & \cos\theta \end{bmatrix} * \begin{bmatrix} \cos\varphi & 0 & \sin(-\varphi) \\ 0 & 1 & 0 \\ \sin\varphi & 0 & \cos\varphi \end{bmatrix} * \begin{bmatrix} Xw \\ Yw \\ Zw \end{bmatrix} \quad \text{Eq. 1}$$

$$Pe = Rz(\Psi) * Rx'(-\theta) * Ry''(-\varphi) * Pw \quad \text{Eq. 2}$$

$$Z - axis = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad \text{Eq. 3}$$

$$X' - axis = \begin{bmatrix} \cos\Psi & \sin\Psi & 0 \\ \sin(-\Psi) & \cos\Psi & 0 \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \cos\Psi \\ -\sin\Psi \\ 0 \end{bmatrix} \quad \text{Eq. 4}$$

$$Y'' - axis = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin(-\theta) \\ 0 & \sin\theta & \cos(\theta) \end{bmatrix} * \begin{bmatrix} \cos\Psi & \sin\Psi & 0 \\ \sin(-\Psi) & \cos\Psi & 0 \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \sin(\Psi) \\ \cos\theta\cos\Psi \\ \sin\theta\cos\Psi \end{bmatrix} \quad \text{Eq. 5}$$

Step 2 – Modeling: The definition of the object coordinate system is determined by the drawing software. The origin is fixed to a pivot point on the object with user-specified geographical location. The geographical location of the world coordinate origin is also given by the GPS measurement; the 3D vector between the object and world coordinate origins can thus be calculated. The methods to calculate the distance between geographical coordinates were originally introduced by Vincenty (1975), and Behzadan and Kamat (2007) proposed a reference point concept to calculate the 3D vector between two geographical locations. This research adopts the same algorithm to place a virtual object in the world coordinate system using the calculated 3D vector. After that, any further translation, rotation, and scaling operations that are needed are applied on the object.

Step 3 and 4 – Viewing Frustum and Projection: The real world is perceived through the perspective projection by the human eye and the web camera. Four parameters are needed to construct a perspective projection matrix: horizontal angle of view, horizontal and vertical aspect ratio, and NEAR plane and FAR plane. All of them together form a viewing frustum and decide the amount of virtual content shown in the augmented space (Figure 6).

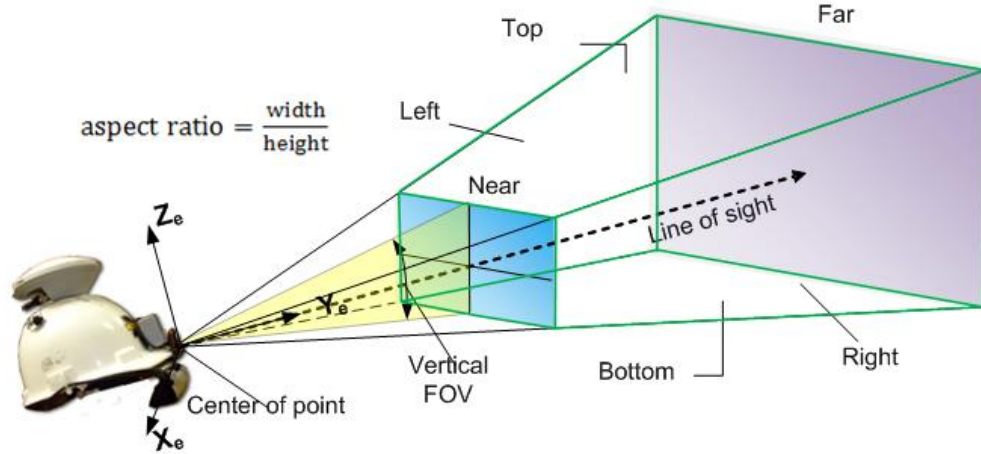


Figure 6 The viewing frustum defines the virtual visible content.

Virtual objects outside of the viewing frustum are either cropped or clipped. It must be noted that although it is not envisioned as a practical visualization aid for an excavator operator, a hard hat-mounted setup shares similar visualization principles which dominates the content included in the viewing volume. This setup was used for algorithm testing and experimentation only. The equipment mounted in the hard hat setup included a camera and a head mounted display in place of a screen monitor.

The NEAR and FAR planes do not affect how the virtual object appears on the projection plane. However, to maintain a high precision z-buffer, the principle is to keep the NEAR plane as far as possible, and the FAR plane as close as possible. The horizontal and vertical angle of view directly influence the magnification of the projected image and are affected by the focal length and aspect ratio of the camera. In order to ensure consistent perspective projection between the real and virtual camera, both of them need to share the same angle of view.

4.1.1 Resolving the Latency Problem in Orientation Sensors

Due to the latency induced by an orientation sensor, correct static registration does not guarantee that the user can see the same correct and stable augmented image when in motion. This section addresses the cause and solution for the dynamic misregistration problem. Dynamic magnetic distortion generally impacts an orientation sensor in motion, and the noise magnification depends on the acceleration of the module. Usually the faster the acceleration is, the higher the noise is (Figure 7). Among the three degrees of freedom, heading is the most sensitive to the noise.

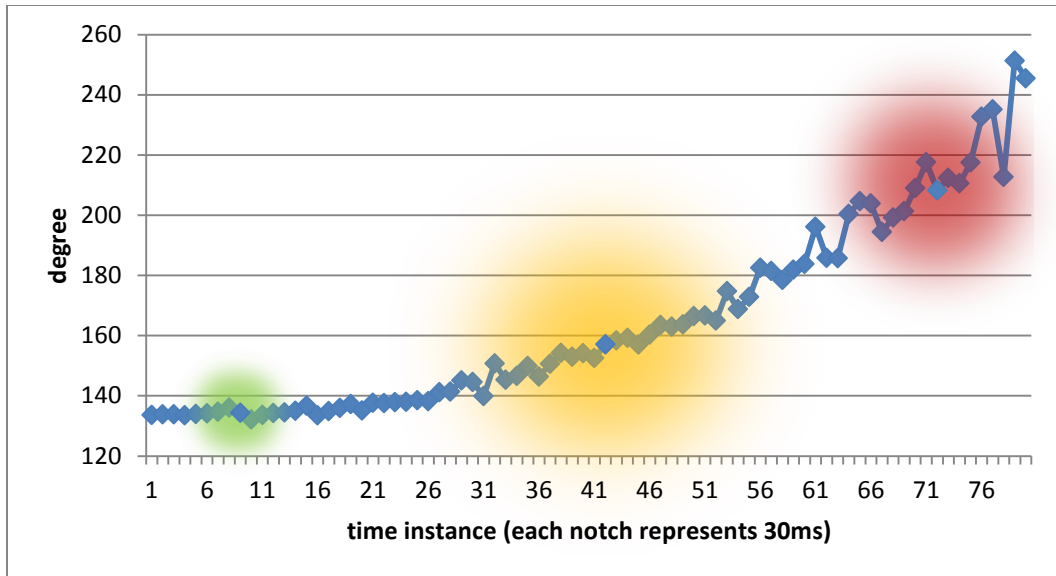


Figure 7. Data show that the noise in the raw data increases as the motion accelerates.

Except the high frequency vibration noise, other types of noise can be removed by a FIR Gaussian filter. A compass module comes with 5 options for filtering: 32, 16, 8, 4, and 0 tap filter. The higher the number is, the more stable the output, but the longer the expected latency. Consider the case of selecting a 32 tap filter (Figure 8). When it is time to send out estimated data at time instant A, the module adds a new sample A to the end of the queue with the first one being dropped, and applies a Gaussian filter to the queue. However, the filtered result actually reflects the estimated value at time instant (A-15). Since the module samples at approximately 30–32 Hz, it induces a 0.5 second delay for a 32 tap filter; a 0.25 second delay for 16 tap filter, and so on. This is called filter-induced latency, and it applies to both PULL and PUSH mode. A 0 tap filter implies no filtering, but significant jittering.

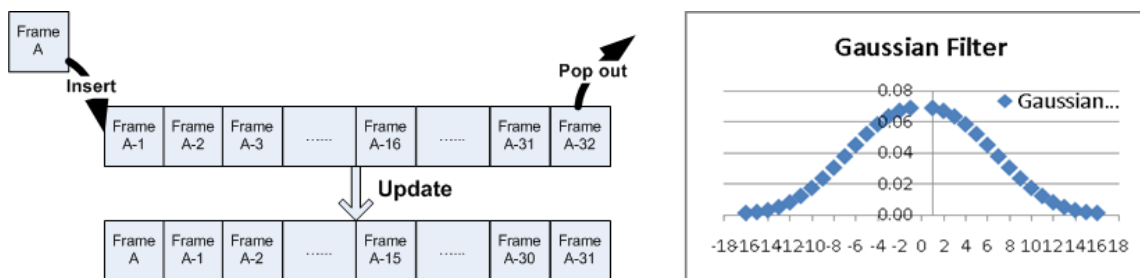


Figure 8. Filter-induced latency when a 32 tap Gaussian filter is used.

In order to avoid the filter-induced latency, the Gaussian FIR filter is moved from the hardware to the software, but with only half window size applied. For example, if a complete Gaussian window is used, it is not until time instant

A+15 that the estimated value can be available for time instant A. However, half window replicates the past data from time instant A-15 to time instant A as the future data from time instant A+1 to A+16, and generates an estimated value for time instant A (Figure 9).

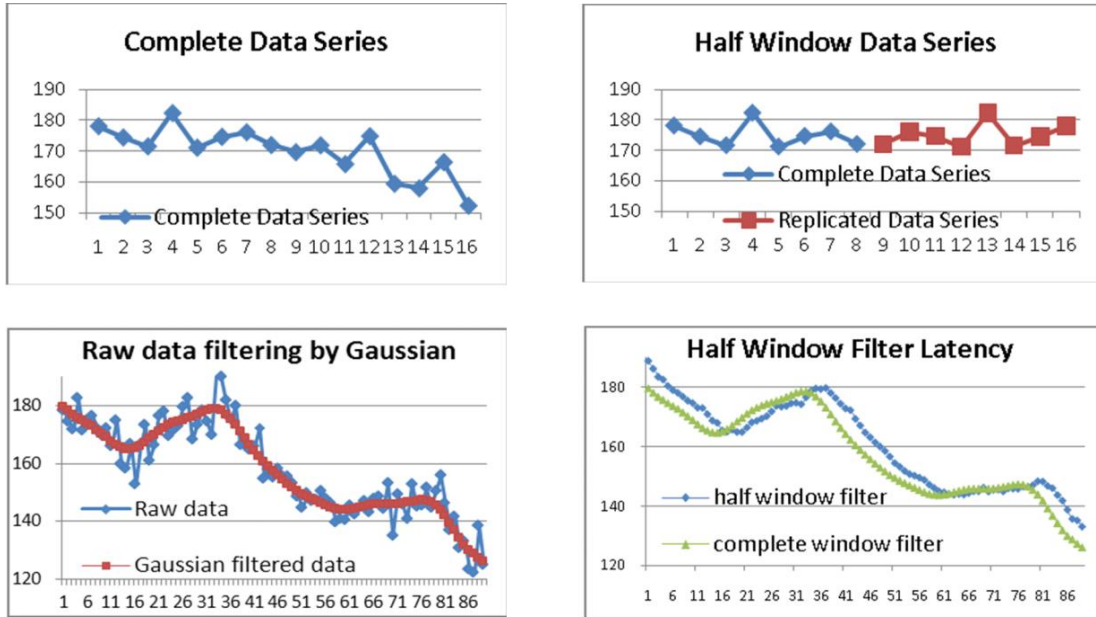


Figure 9. Half-window filter latency (each notch on the X-axis represents 30 ms).

Nevertheless, as is shown in the graph chart, half window still causes 4–5 frames of latency on average. Depending on the speed of module movement, the faster the speed, the longer latency it presents. The project team addresses this kind of latency as half window induced latency. Because the half window Gaussian filter puts more emphasis on the current frame, it makes the estimated result more sensitive to noise contained in the current frame, and consequently there is more jittering than in the estimated result of the complete window Gaussian filter. Therefore, a second half window Gaussian is applied on the first filtered result for smoothing purposes, but this introduces 1–2 extra frames of latency (Figure 10).

However, this additional latency can be discounted because it does not exceed the original latency—the one between the half window Gaussian filter and the complete window Gaussian filter. Therefore, double the additional latency is subtracted from the Twice Gaussian filter result, and this makes the estimation closer to the actual data than the half window Gaussian filter result. Unfortunately, this approach fails during the transition state, and leads to overshooting during change of direction, and during transitions from dynamic to static states. It must be noted that acceptable latency is a very subjective quantity that could not be explicitly measured. In Augmented Reality applications, the amount of acceptable latency is dependent on a user's tolerance to motion cues, and this will vary from user to user and their reaction to motion cues.

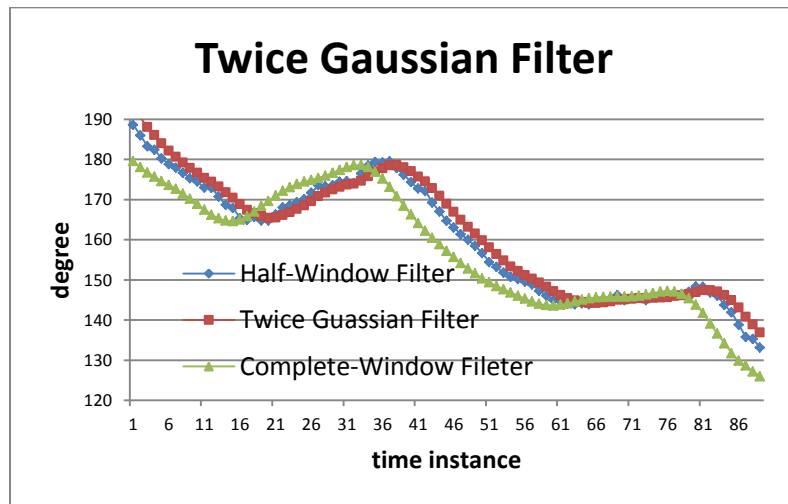
4.1.2 Adaptive Latency Compensation Algorithm

In order to resolve the overshooting problem, the estimated result needs to be forced to the observed data when the module comes to a stop. This is possible because the observed data is very stable and close to the actual value when the module is static. Large collections of observed value show standard deviation is a good indicator of dynamic and static status of the sensor; when the standard deviation is larger than 6, the heading component of the module is in motion, otherwise it is in static or on the way to stopping (Figure 11).

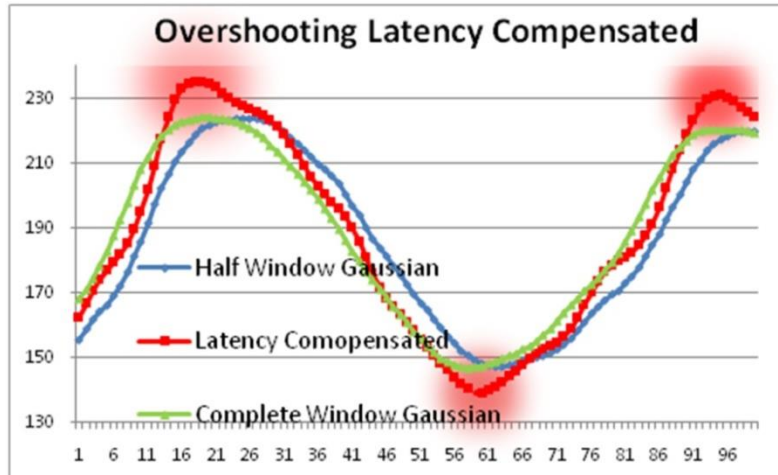
Therefore the adaptive algorithm computes the latency compensated value as follows: when the standard deviation is no larger than 6, the compensated value is double of the subtraction of the Twice Gaussian filter by the half Window Gaussian filter result; otherwise the compensated value is equal to the subtraction of the Twice Gaussian Filter by the observed data.

4.1.3 Field Demonstration of Augmented Reality Registration Algorithms

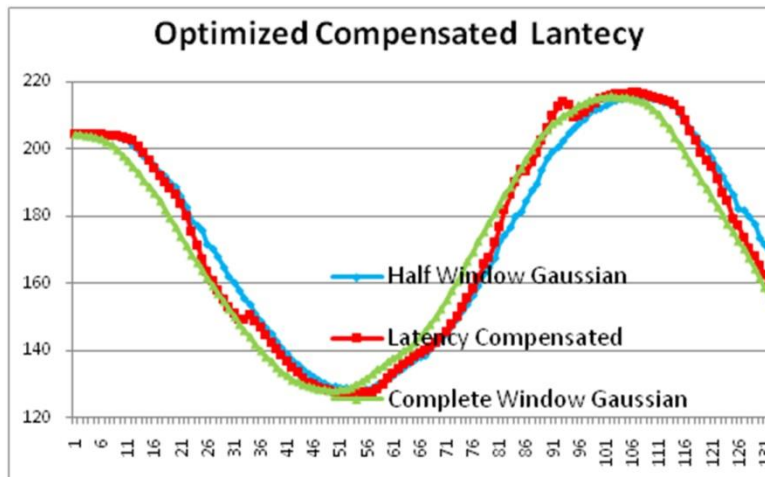
The robustness of the developed registration algorithms was demonstrated in field experiments. Electricity conduits in the vicinity of the G.G. Brown Building at the University of Michigan were exported as KML (Keyhole Modeling Language) files from a Geodatabase provided by the DTE Energy Company. The following procedure interprets KML files and builds conduit models (Figure 12):



(A) Additional Latency



(B) Overshooting Problem



(C) Adaptive latency compensation

Figure 10. Adaptive latency compensation algorithm (each notch on the X-axis represents 30 ms).

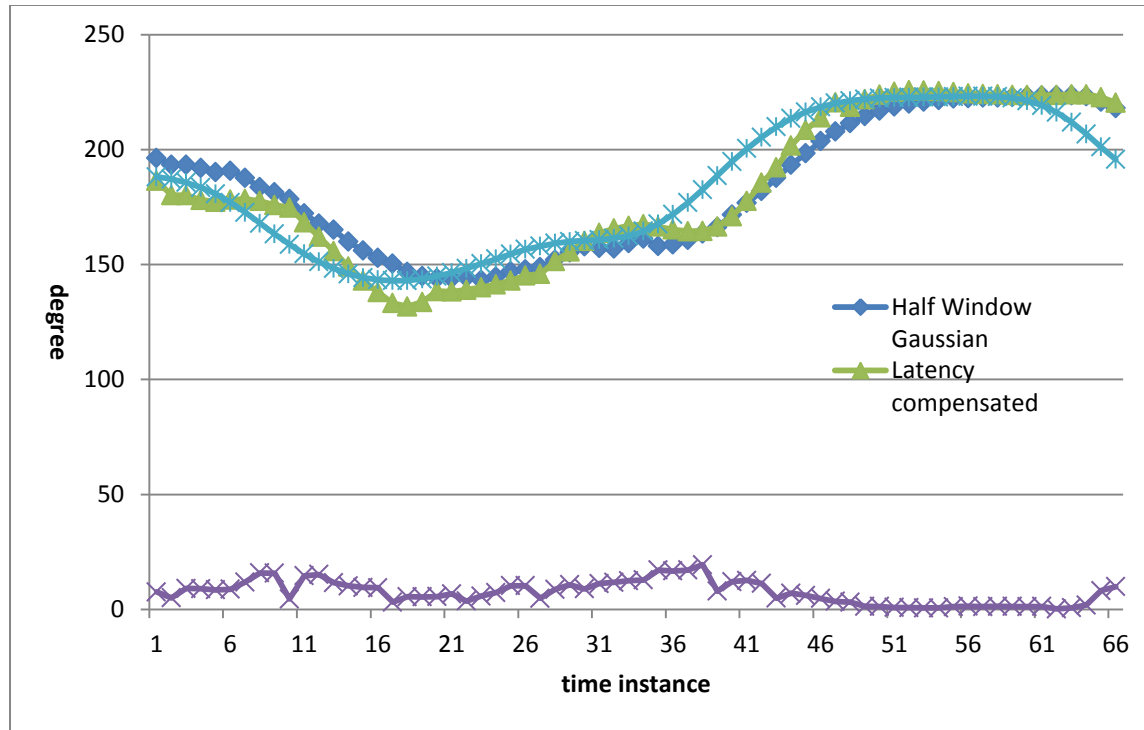


Figure 11. Standard deviation indicates the motion pattern (each notch on the X-axis represents 30 ms).

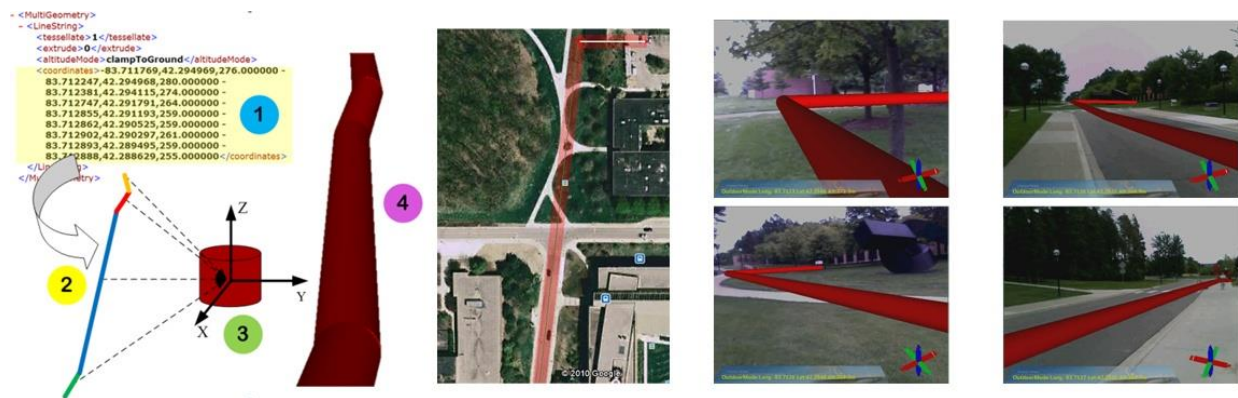


Figure 12. Conduit loading procedure, conduits overlaid on Google Earth and field experiment results.

- (1) Extract the spatial and attribute information of pipelines from the KML file using libkml, a library for parsing, generating, and operating in KML (Google 2008). For example, the geographical location of pipelines is recorded under the Geometry element as “LineString” (Google 2012). A cursor is thus designed to iterate through the KML file, locate “LineString” elements, and extract the geographical locations.

- (2) Convert consecutive vertices within one “LineString” from the geographical coordinate to the local coordinate in order to raise computational efficiency during the registration routine. The first vertex on the line string is chosen as the origin of the local coordinate system, and the local coordinates of the remaining vertices are determined by calculating the relative 3D vector between the rest of the vertices and the first one, using the Vincenty algorithm.
- (3) In order to save storage memory, a unit cylinder is shared by all pipeline segments as primitive geometry upon which the transformation matrix is built.
- (4) Scale, rotate, and translate the primitive cylinder to the correct size, attitude, and position. For simplicity, the normalized vector between two successive vertices is named as the pipeline vector. First, the primitive cylinder is scaled along the X- and Y-axis by the radius of the true pipeline, and then scaled along the Z-axis by the distance between two successive vertices. Second, the scaled cylinder is rotated along the axis—formed by the cross product between vector $\langle 0, 0, 1 \rangle$ and the pipeline vector—by the angle of the dot product between vector $\langle 0, 0, 1 \rangle$ and the pipeline vector. Finally, the center of the rotated pipeline is translated to the midpoint between two successive vertices. This step is applied to each pair of two successive vertices.

It must be noted that Google Earth is only used for experimentation and demonstration of electric conduit data conversion. The underlying algorithms themselves are not tied to the use of Google Earth and are compatible with any geospatial data system. Dynamic misregistration continues to be an open research problem and is still under investigation. Several efforts are being made: (1) synchronizing the captured image and sensor measurements; and (2) optimizing the adaptive latency compensation algorithm with image processing techniques (e.g., optical flow can afford more information about the angular speed). Results obtained from the conducted field experiments to evaluate the developed algorithms are shown in the following Figures 13 and 14.

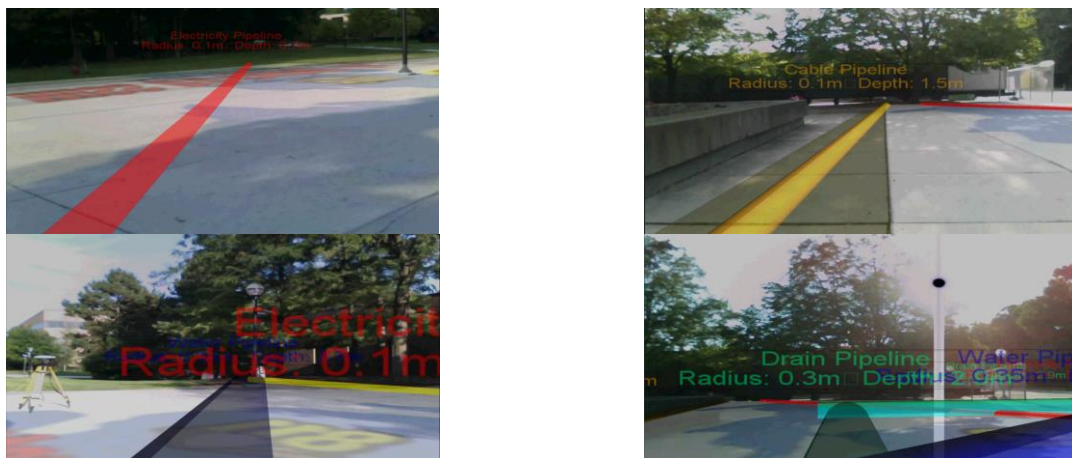


Figure 13. Labeling attribute information and color coding on the underground utilities.

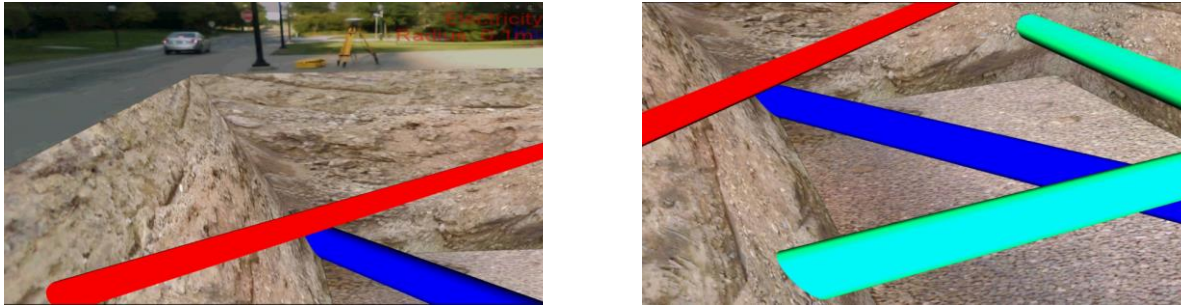


Figure 14. X-ray view of the underground utilities.

Dynamic misregistration primarily pertains to the human perception of the AR graphics' authenticity. Given the typical speed of an excavator when it is digging, the project team expects that most human users' tolerance for dynamic misregistration can overcome this issue during visualization.

4.2 REAL-TIME POSE ESTIMATION OF ARTICULATED EXCAVATOR

The pose of an articulated machine includes the position and orientation of not only the machine base (e.g., tracks or wheels), but also each of its major articulated components (e.g., stick and bucket). The ability to automatically estimate this pose is a crucial component of improving an operator's spatial awareness to underground utilities. A computer vision-based solution using a network of cameras and markers was developed in this research to enable such a capability for articulated machines (Figure 15).

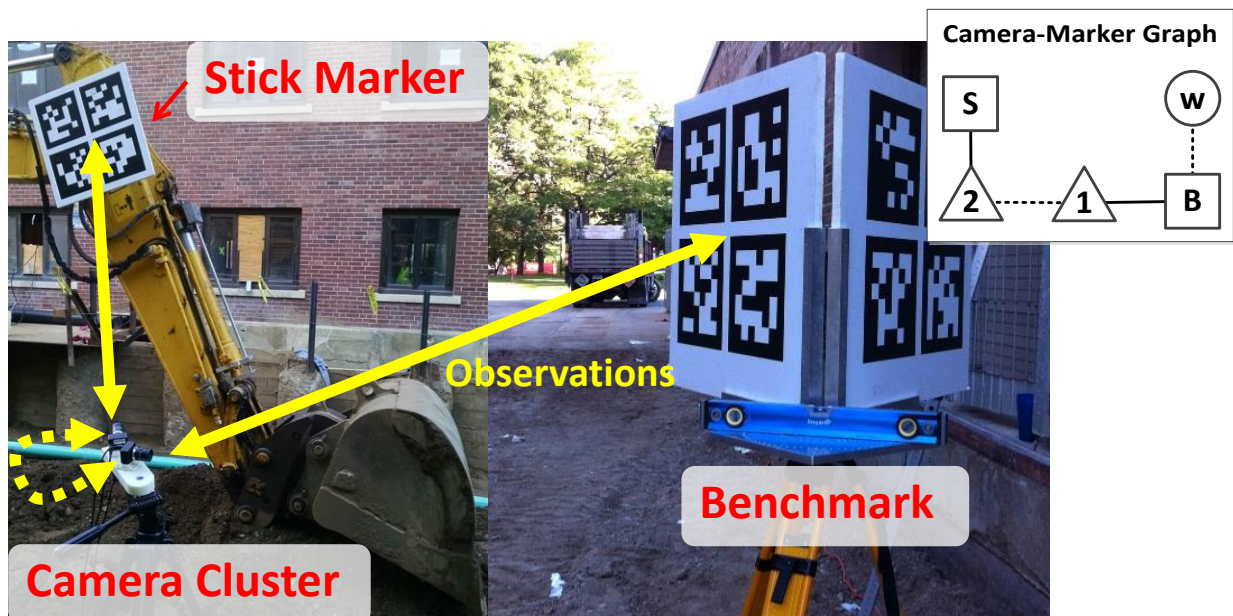


Figure 15. Overview of computer vision-based pose estimation system.

Different versions of the proposed articulated machine pose estimation system design are explained first. Then, the process to calibrate this system is described. Finally, uncertainty analysis is described for the system with some important observations of the relationship between the system configuration and its stability; that is, uncertainty of the estimated excavator pose for AR visualization.

The designed computer vision-based articulated machine pose estimation solution relies on a method called marker-based pose estimation. Generally, marker-based pose estimation firstly finds a set of 2D geometry features (e.g., points or lines) on an image captured by a calibrated camera, then establishes correspondences between another set of 2D or 3D geometry features on a marker whose pose is known with respect to a certain coordinate frame of interest, and finally estimates the pose of the camera in that coordinate system. If 2D-2D correspondences are used, the pose is typically estimated by homography decomposition. If 2D-3D, the pose is typically estimated by solving the perspective-n-point (PnP) problem. Two typical marker-based pose estimation methods are AprilTag (Olson, 2011) and KEG (Feng and Kamat 2012) algorithms.

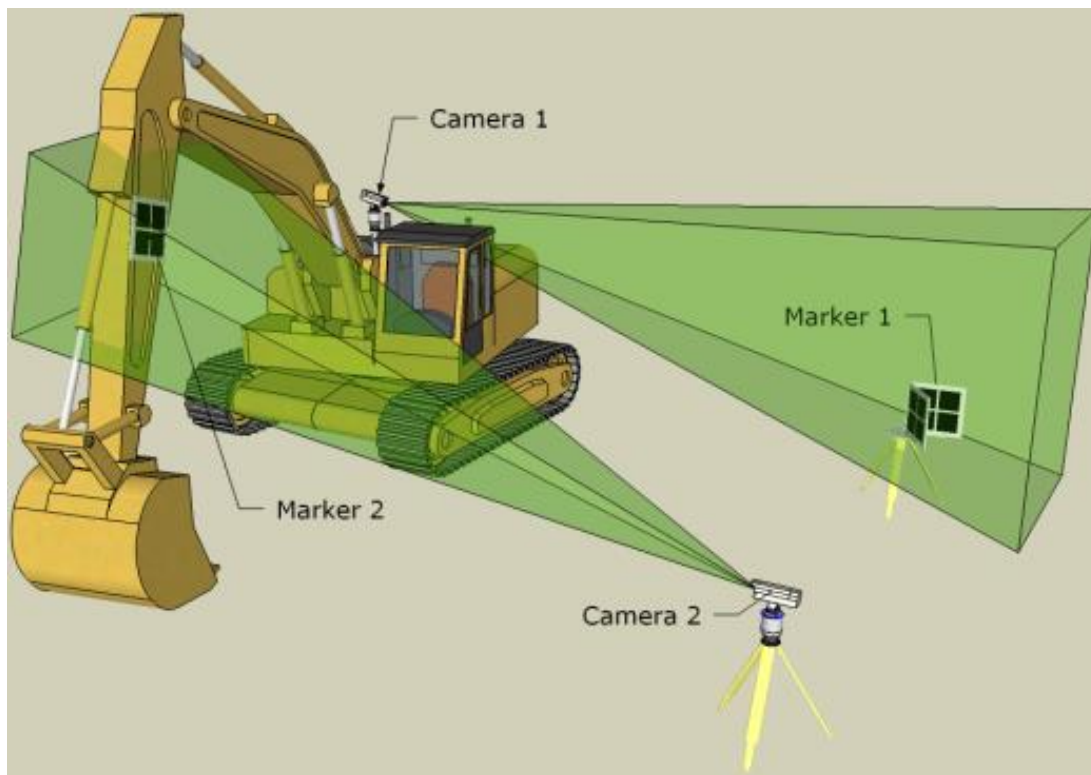


Figure 16. Two examples of basic camera marker configuration.

There are two ways of applying marker-based pose estimation for poses of general objects of interest. As shown in Figure 16, one way is to install the calibrated camera 1 rigidly on the object of interest (in this case, the cabin of the

17

excavator), and pre-survey the marker 1's pose in the project coordinate frame. The other way is to install the marker 2 rigidly on the object (in this case, the stick of the excavator), and pre-calibrate the camera 2's pose in the project coordinate frame. As long as the camera 2 (or the marker 1) stays static in the project coordinate frame, the pose of the excavator's stick (or the cabin) can be estimated in real-time.

However, these basic configurations do not always satisfy application requirements. For example, if only the camera 1 and the marker 1 are used, the excavator's stick pose cannot be estimated. On the other hand when only the camera 2 and the marker 2 are used, once the stick leaves the field of view (FOV) of the camera 2, the stick's pose becomes unavailable as well. Thus it is necessary to take a camera's FOV into consideration when designing an articulated machine pose estimation system. This understanding leads to the camera marker network design proposed.

4.2.1 Camera Marker Network

A camera marker network is an observation system containing multiple cameras or markers for estimating poses of objects embedded in this system. It can be abstracted as a graph with three types of nodes and two types of edges (e.g., Figure 17). A node denotes an object pose (i.e., the local coordinate frame of that object), which can be a camera, a marker, or the world coordinate frame. An edge denotes the relative relationship between two objects connected by this edge, which can be either image point observations for the previously mentioned marker-based pose estimation, or a known pose constraint (e.g., through calibration).

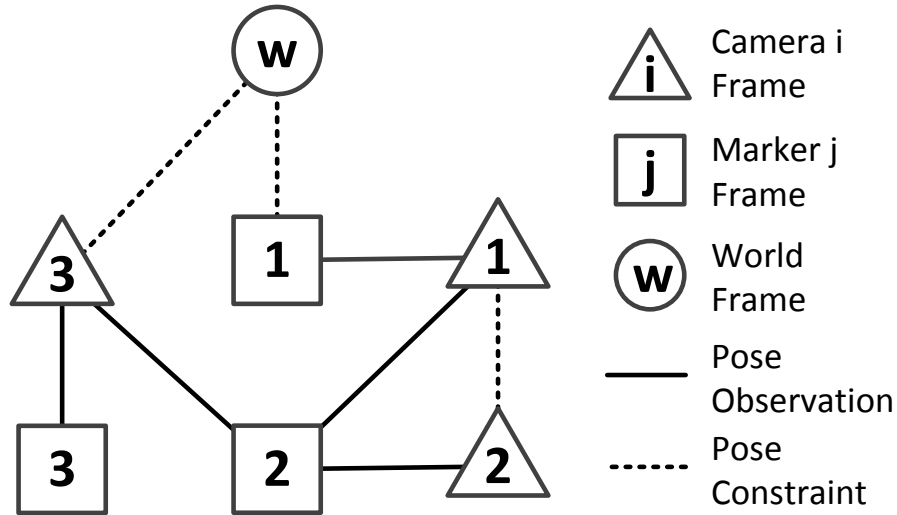


Figure 17. An example graph of a camera marker network.

Thus, if at least one path exists between any two nodes in such a graph, the relative pose between them can be estimated. In addition, any loop in the graph means a constraint of poses that can be used to improve the pose estimation. For example, in Figure 17, marker 2's pose in the world frame can be found via a path through camera 3

whose own pose in the world frame is pre-calibrated. The marker 2's pose can also be better estimated when observed by the rigidly connected camera 1 and 2 whose relative pose is pre-calibrated, since a loop is created.

Applying this concept to articulated machine pose estimation results in numerous possible designs. One of the possible camera marker networks is shown in Figure 18, camera 1 observes the benchmark while camera 2 observes the stick marker, and the rigid transformation between the two cameras is pre-calibrated. Thus as long as the two markers stay inside the two cameras' FOV respectively, the stick's pose in the world frame can be estimated.

It is worth noting that this only illustrates a simple configuration. With more cameras and markers in the network, there are more chances of creating loops and thus improving pose estimation, especially considering that surveillance cameras are becoming popular in construction job sites whose poses can be pre-calibrated and thus act as the camera 3 in Figure 17.

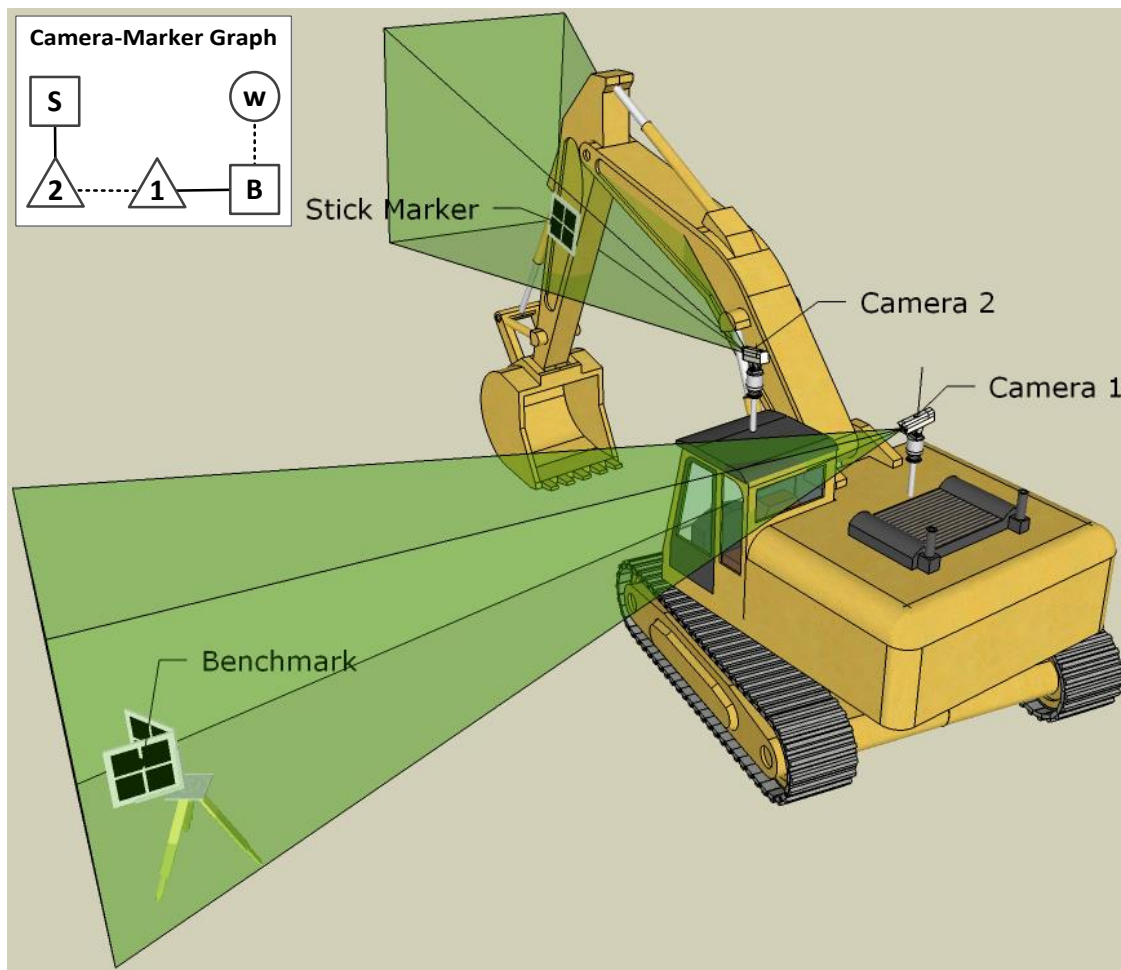


Figure 18. Multiple-camera multiple-marker configuration.

4.2.2 Experimental Prototypes

Multiple prototypes have been implemented to realize the above described camera marker network designs. Figure 19 demonstrates one of the early prototypes implementing a single-camera multiple marker configuration. A mechanical device driven by a synchronous belt was adopted to map the relative rotation between the excavator bucket and the stick to the relative rotation between the stick marker and the rotary marker. This implementation enables pose tracking of the excavator bucket.

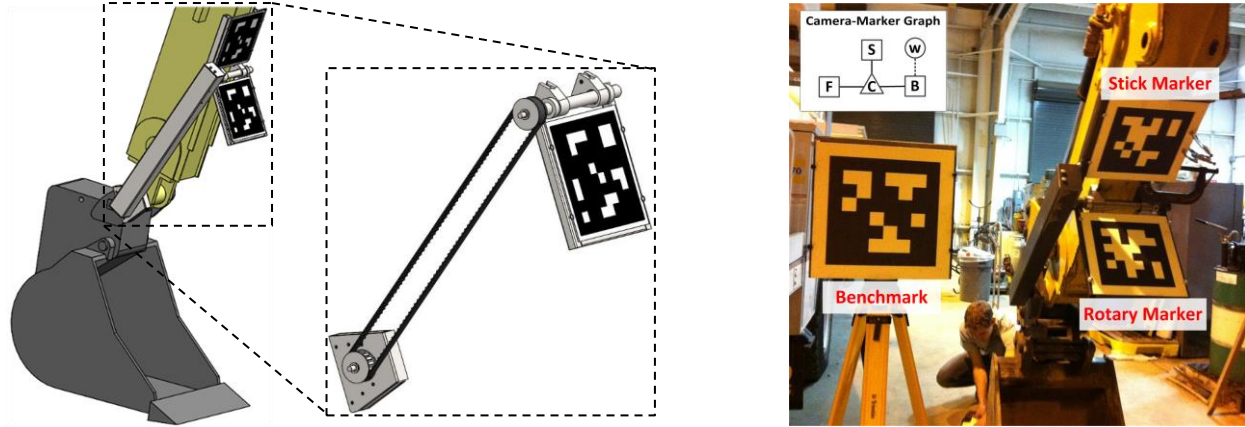


Figure 19. Synchronous belt prototype design.

The prototype functions by means of two markers. The first marker, termed stick marker, is rigidly attached to the stick with a known relationship to the bucket's axis of rotation. The second marker, termed rotary marker, is attached at a location removed from the vicinity of the bucket. The rotary marker is constrained with one degree of rotational freedom and a known angular relationship to the bucket. If the bucket's geometry is also known, or measured onsite, then all necessary information is available to deduce tooth pose, as shown in Figure 20.

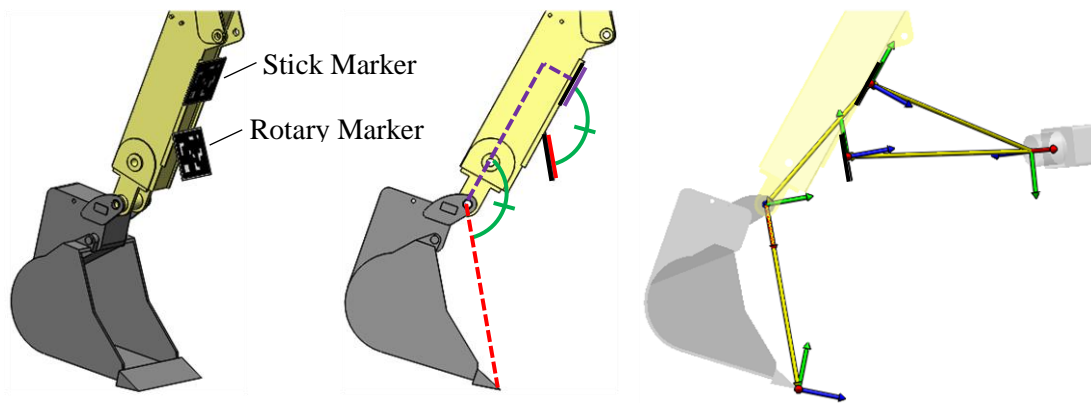


Figure 20. Bucket tooth pose calculation in the synchronous belt prototype.

Due to the potential interference of the rotary marker and any obstructions during excavation, the above synchronous belt prototype was slightly modified and evolved to the current prototype as shown in Figure 21. The newer working prototype implements the multiple-camera multiple-marker configuration similar to Figure 19. Two cameras are rigidly mounted forming a camera cluster. A cable potentiometer is installed on the bucket's hydraulic cylinder to track the relative motion of the excavator bucket and the stick even if the bucket is deep inside the earth. In addition to possessing a cable potentiometer for measuring linear displacement, the device contains a microcontroller for signal conversion, a radio for wireless transmission, and a battery for power, all of which are mounted inside an enclosure for protection. In the project team's experiments, the battery has been shown to last for up to 3 days before it has to be re-charged. No additional signal interference has been observed in the sensor suite either.

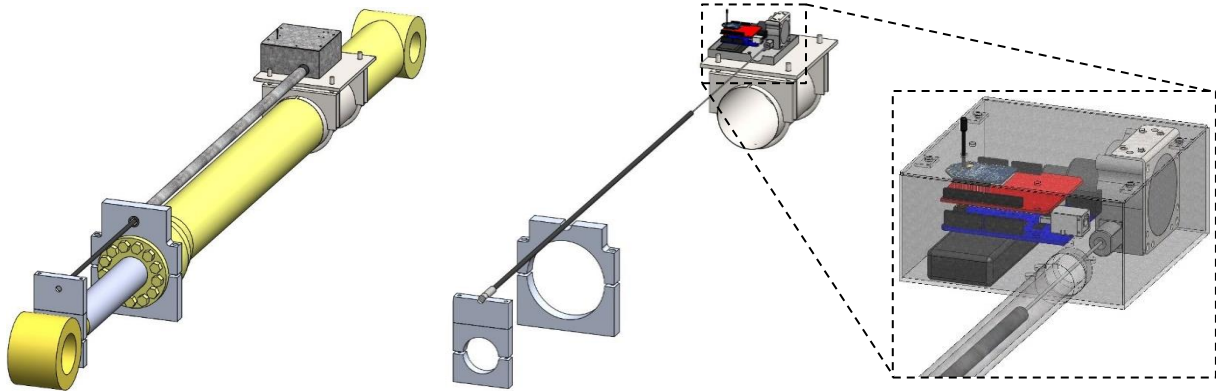


Figure 21. Cable potentiometer prototype.

4.2.3 System Calibration

Three types of calibration are necessary for an articulated machine pose estimation system implementing the above camera marker network design. The first type is intrinsic calibration which determines internal parameters (e.g., focal length) of all cameras in the system. This is done using same methods as in (Feng et al. 2014). The second type is extrinsic calibration which determines relative poses (e.g., dotted edges in the graph) designed to be calibrated before system operation. There are two kinds of such poses: (1) poses of static markers in the world frame, and (2) poses between rigidly connected cameras. The first kind of poses can be calibrated by traditional surveying methods using a total station. The second kind of poses, however, cannot be directly surveyed physically since a camera frame's origin and principal directions usually cannot be found or marked tangibly on that camera.

Thus to calibrate a set of m rigidly connected cameras, a camera marker graph needs to be constructed as denoted in Figure 22. A set of n markers' poses need to be surveyed in the world frame. Then when the m cameras observe these n calibration markers, the graph is formed to estimate each camera's pose in the world frame and thus their

relative poses between each other (i.e., edges with question mark) are calibrated. It is suggested to ensure that multiple loops exist in this graph to improve the accuracy of the poses to be calibrated. Such a loop exists as long as at least two markers are observed by the same camera simultaneously. It is also worth noting that with enough calibration markers, each camera's intrinsic parameters can be further optimized with their extrinsic parameters.

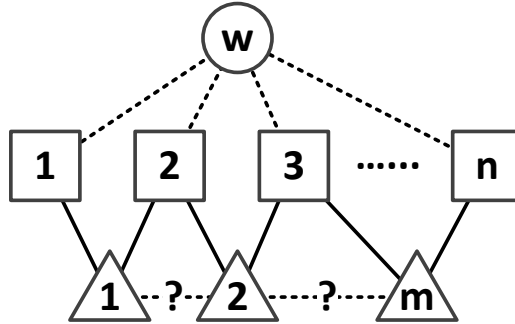


Figure 22. Camera marker graph for extrinsic calibration.

The third and final type of calibration involves mapping a relationship between the cable potentiometer output and the pose of the excavator's bucket. The calibration process requires a marker to be attached to the side of the bucket for the duration of the calibration process. The system is then calibrated by moving the bucket through its range of motion and mapping the potentiometer output to the relative transformation between the stick marker and bucket teeth. After the calibration process, the marker is removed from the bucket. The system then functions by measuring stick marker pose and cylinder stroke length, and mapping such measurements to bucket tooth pose. The marker on the bucket can be calibrated before operation and at off-site locations if desired. Once calibrated, the relationship between the bucket rotation and potentiometer extension can be saved to a file. Multiple buckets can be calibrated off-site and saved in this manner.

It must be noted that the overall purpose of the calibration is to characterize the change of bucket teeth elevation due to the bucket rotation. This information can be relayed to the cable length change of a potentiometer. The whole calibration process is automated and the operator only needs to slowly rotate the bucket in a full cycle while the camera can cover both markers. The calibration only needs to be done once, and this takes about 5 minutes. Once the site benchmark is set up, it maintains the achieved accuracy up to 100 ft. When an excavator is digging around utilities and hand digging is involved, 100 ft is a reasonable projected daily progress. However, if an excavator does move beyond 100ft, it may be possible to include multiple site benchmarks. The additional benchmarks can likely be calibrated relative to the initial using computer vision methods such as structure from motion. However, confirming this possibility requires additional research and experimentation.

4.2.4 Uncertainty Analysis

It is not sufficient to only estimate the pose of an articulated machine. The uncertainty of the estimated pose is critical for the following reasons. Firstly the uncertainty provides a measure of the confidence level of the estimated pose, which is necessary for many downstream applications (e.g., deciding buffer size for collision avoidance). Secondly it serves as a tool for evaluating the stability of the pose estimation system under different configurations, and thus further guiding to avoid critical configurations that lead to unstable pose estimation.

To perform uncertainty analysis on the proposed camera marker network pose estimation system, the system is firstly abstracted as the following state space model:

$$\mathbf{Z} = \mathbf{F}(\mathbf{X}; \mathbf{Y}, \mathbf{C}) \quad \text{Eq. 6}$$

where \mathbf{X} is the state vector of the network (usually encodes the poses of nodes in the graph), \mathbf{Z} is the predicted measurement vector containing image coordinates of all the points projected from markers, \mathbf{Y} is the known parameters (usually contains marker points' local coordinates), \mathbf{C} is the calibrated parameters (usually encodes all cameras' intrinsic parameters and all calibrated poses), and \mathbf{F} is the system's observation function parameterized by \mathbf{Y} and \mathbf{C} ; that is, the camera perspective projection function.

For example, for a network of a single camera and a single marker, \mathbf{X} is a 6 x 1 vector that encodes the marker's orientation and position in the camera frame; \mathbf{Y} is a $3n \times 1$ vector containing n marker points' coordinates from surveying; \mathbf{C} is a vector of the camera intrinsic parameters. If another marker is added to this network, \mathbf{Y} should be extended with points on the new marker.

4.2.5 Uncertainty Propagation

No matter how complex such a network is and what method is used to get an initial estimate of \mathbf{X} (either PnP or homograph decomposition), the optimized state $\hat{\mathbf{X}}$ can be calculated by the following least square optimization; that is, bundle adjustment:

$$\hat{\mathbf{X}} = \arg \min_{\mathbf{X}} \left\| \hat{\mathbf{Z}} - \mathbf{F}(\mathbf{X}; \mathbf{Y}, \mathbf{C}) \right\|_{\mathbf{P}_Z}^2 \quad \text{Eq. 7}$$

where \mathbf{P}_Z is the a priori covariance matrix of the actual measurements $\hat{\mathbf{Z}}$, typically assumed as $\sigma_u^2 \mathbf{I}$ when image coordinates are measured with a standard deviation of σ_u .

To backward propagate the measurement uncertainty \mathbf{P}_Z to the uncertainty of the optimized state $\hat{\mathbf{X}}$ requires linearization of \mathbf{F} around $\hat{\mathbf{X}}$. Since the error is assumed to come from only the measurements (the uncertainty in calibrated parameters \mathbf{C} can be included in future work, but is assumed to be negligible in this paper), one can directly apply the results in (Hartley and Zisserman 2000) to calculate the uncertainty of the optimized states:

$$\mathbf{P}_{\hat{\mathbf{X}}} = (\mathbf{J}^T \mathbf{P}_Z^{-1} \mathbf{J})^{-1} = \sigma_u^2 (\mathbf{J}^T \mathbf{J})^{-1} \quad \text{Eq. 8}$$

where $\mathbf{J} = \left. \frac{\partial \mathbf{F}}{\partial \mathbf{X}} \right|_{\hat{\mathbf{X}}}$ is the Jacobian matrix of \mathbf{F} evaluated at $\hat{\mathbf{X}}$.

4.2.6 Uncertainty and Configuration

Equation (3) not only provides a means of evaluating uncertainty of the optimized pose estimation of a camera marker network, but also provides a tool to predict the system stability at any given system configuration \mathbf{X} before even making any measurements. This is done by evaluating the Jacobian matrix \mathbf{J} of \mathbf{F} at that \mathbf{X} , and then applying equation (3) to predict the covariance matrix. It is based on the fact that the aforementioned backward propagation of measurement uncertainty does not directly rely on specific measurements. In fact it directly relies on the system configuration \mathbf{X} around which the linearization is performed. Thus, when evaluating Jacobian matrix \mathbf{J} at a configuration \mathbf{X} , equation (3) yields the theoretically best/smallest pose estimation uncertainty one can expect at that configuration, which denotes the system stability at that configuration.

Using this method, some important empirical conclusions on the basic single-camera single-marker system are found about relationships between system stability and configuration, based on numerical experiments, which are useful for more complex system design and are listed as follows. Similar analysis can be performed to multiple-camera or multiple-marker system in future work.

- (1) The marker's origin/position in the camera frame, ${}^c\mathbf{t}_m$, has the largest uncertainty along a direction nearly parallel to the camera's line of sight to the marker; that is, ${}^c\mathbf{t}_m$ itself. Figure 23 exemplifies this observation at two randomly generated poses between the camera and the marker.
- (2) The largest uncertainty of the marker's position in the camera frame increases approximately quadratic to the marker's distance to the camera, compared to which the two smallest uncertainty's increases are almost negligible. Figure 24 shows a typical example.
- (3) The largest uncertainty of the marker's position in the camera frame increases approximately linear to the camera focal length, compared to which the two smallest uncertainty's increases are almost negligible. Figure 25 shows a typical example.

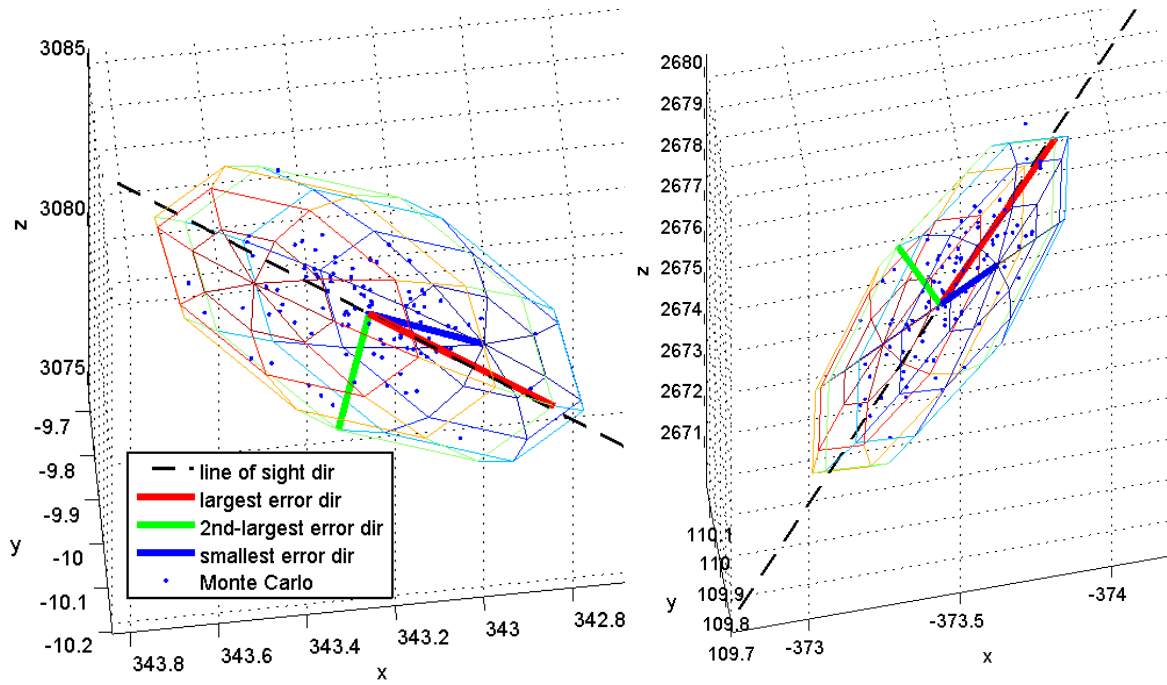


Figure 23. Largest position error direction.

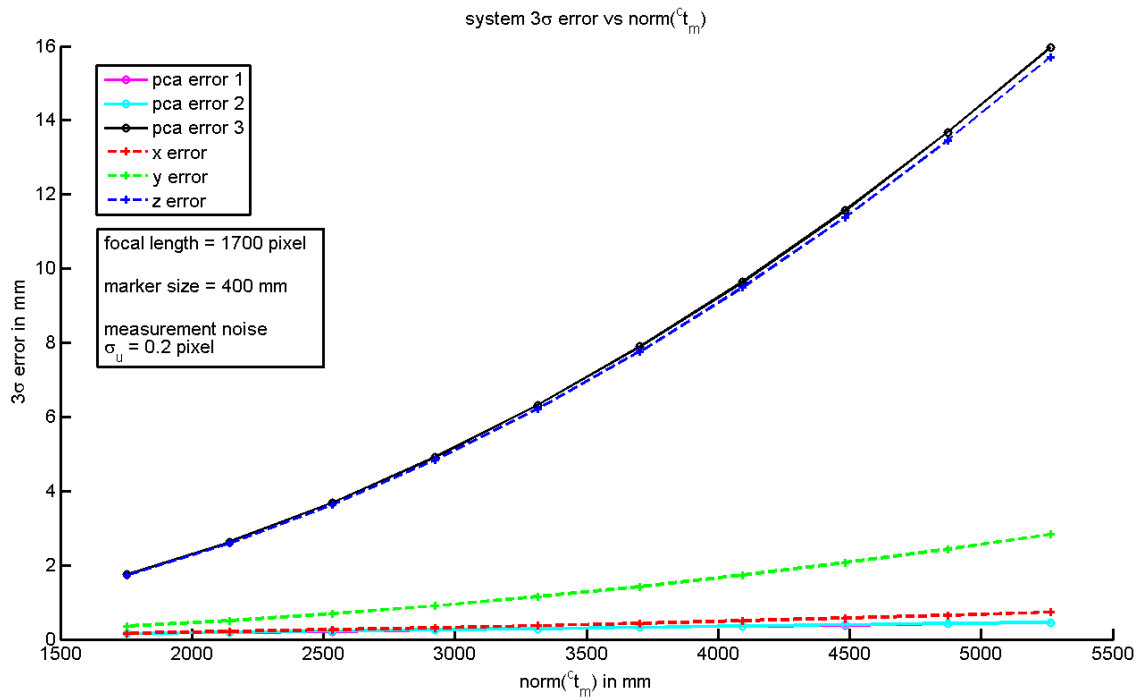


Figure 24. Position error vs. marker distance.

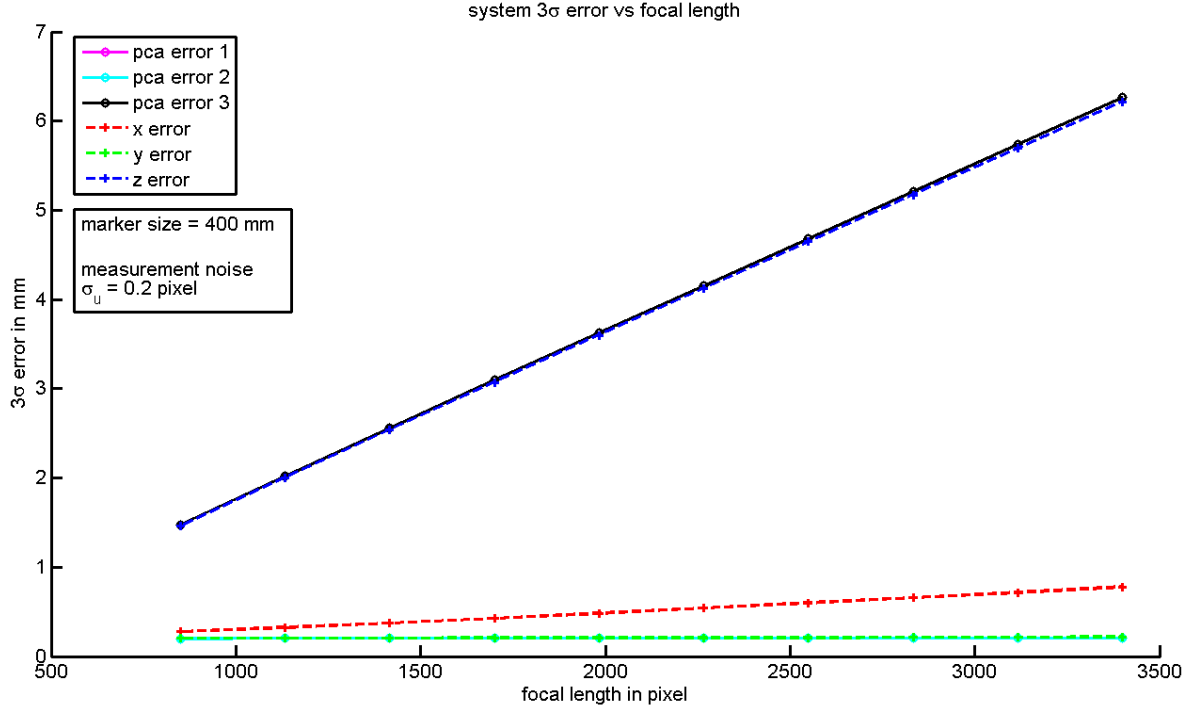


Figure 25. Position error vs. focal length.

4.2.7 Feasibility Experiments

Before implementing the pose estimation system prototypes, a set of experiments were performed to test the feasibility of marker-based pose estimation in different indoor/outdoor construction environments. In all the experiments, AprilTag (Olson 2011) was chosen as the basic marker detection and tracking algorithm. Firstly, the outdoor detectability of markers was tested. A marker's detectability is a function of many factors including the marker size, the distance between the marker and the camera, included angle between the camera viewing direction and the marker plane's normal direction, and also image resolution. Since the distance between the marker and the camera is the most critical factor affecting the method's feasibility in real applications, this experiment is performed by fixing other factors and then gradually increasing the distance of the marker in front of the camera, until the algorithm fails to detect the marker, and recording the distance. Varying other factors and repeating this process results in Table 2. One can consult this table to decide how large the marker should be to fit application need.

Table 2. Outdoor Detectability of AprilTag

Max Detectable Distance (m)		Marker Angle (degree)			
		0	45	0	45
Marker Size (m ²)	0.2 x 0.2	6.10	4.88	11.28	8.84
	0.3 x 0.3	8.23	7.01	14.94	11.58
	0.46 x 0.46	13.41	11.28	25.91	21.64
	0.6 x 0.6	19.51	16.46	34.44	30.48
Image Resolution		640 x 480		1280 x 960	
Focal Length		850 pixels		1731 pixels	
Processing Rate		20 Hz		5 Hz	

Second, illumination is a critical factor affecting performance of many computer vision algorithms. The AprilTag algorithm was thus tested under various illumination conditions to examine its robustness for construction applications. Figure 26 shows successful marker detection under different indoor/outdoor lighting conditions. These experiments and following extensive prototype tests proved the AprilTag-based marker detection method's robustness to illumination changes.

In subsequent outdoor experiments described ahead, the cameras detect markers very well even in overcast weather. In fact, most of the time, it is necessary to shrink the aperture of the camera so that the image is not over exposed. However, in order to deal with complicated illumination conditions in the outdoor environment, techniques for the automatic exposure adjustment in the sun have also been designed.

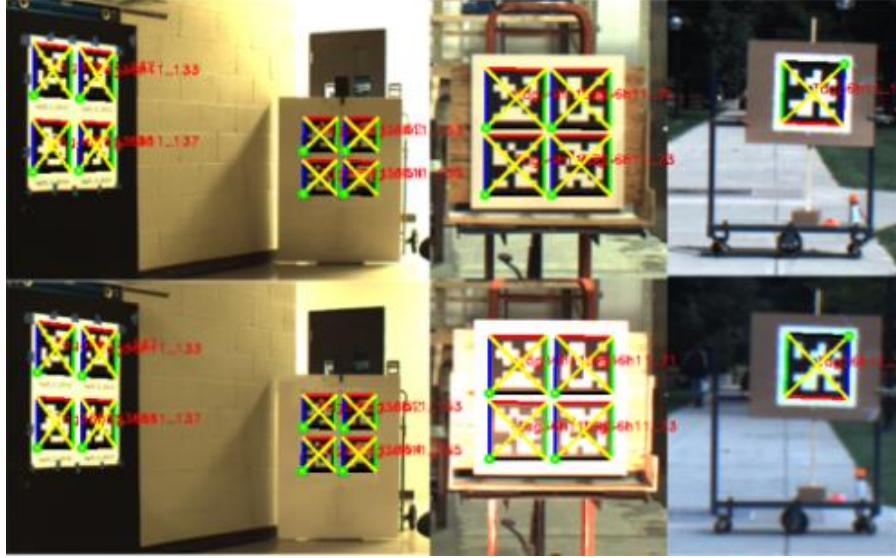


Figure 26. Marker detection vs. illumination.

Finally, for uncertainty propagation, one needs to have a prior estimation of the image measurement noise's standard deviation σ_u . This is achieved by collecting multiple images under a static camera marker pose. Repeating this process for different poses and collecting corresponding image measurement statistics lead to an image measurement covariance matrix Σ_u , which can be further relaxed to $\sigma_u^2 \mathbf{I}$ to include all the data points. Figure 27 shows that $\sigma_u = 0.2$ pixel is reasonable.

As previously mentioned, a multiple-camera multiple-marker articulated machine pose estimation prototype has been implemented with the application of estimating an excavator's bucket depth in a project frame, which could be used for automatic excavation guidance and grade control.

The top row of Figure 28 (a) shows the camera cluster of the prototype in Figure 16, and different experiment configurations to test the depth estimate's accuracy. The experiments were set up by observing the two markers in the bottom row of Figure 28 (a) using the two cameras in the cluster respectively. Then the depth difference between the two markers was estimated using the proposed method, while the ground truth depth difference between the two marker centers was measured by a total station with 1 mm accuracy. Figure 28 (b) illustrates the configurations of different sets of such experiments, for comprehensive tests of the method's accuracy under several system and design variations. The first set varies one of the marker's pitch angle (top row of the figure). The second set varies its height (bottom-left). The third set varies its distance to the camera (bottom-middle). And the fourth set varies the number of tags used in that marker (bottom-right).

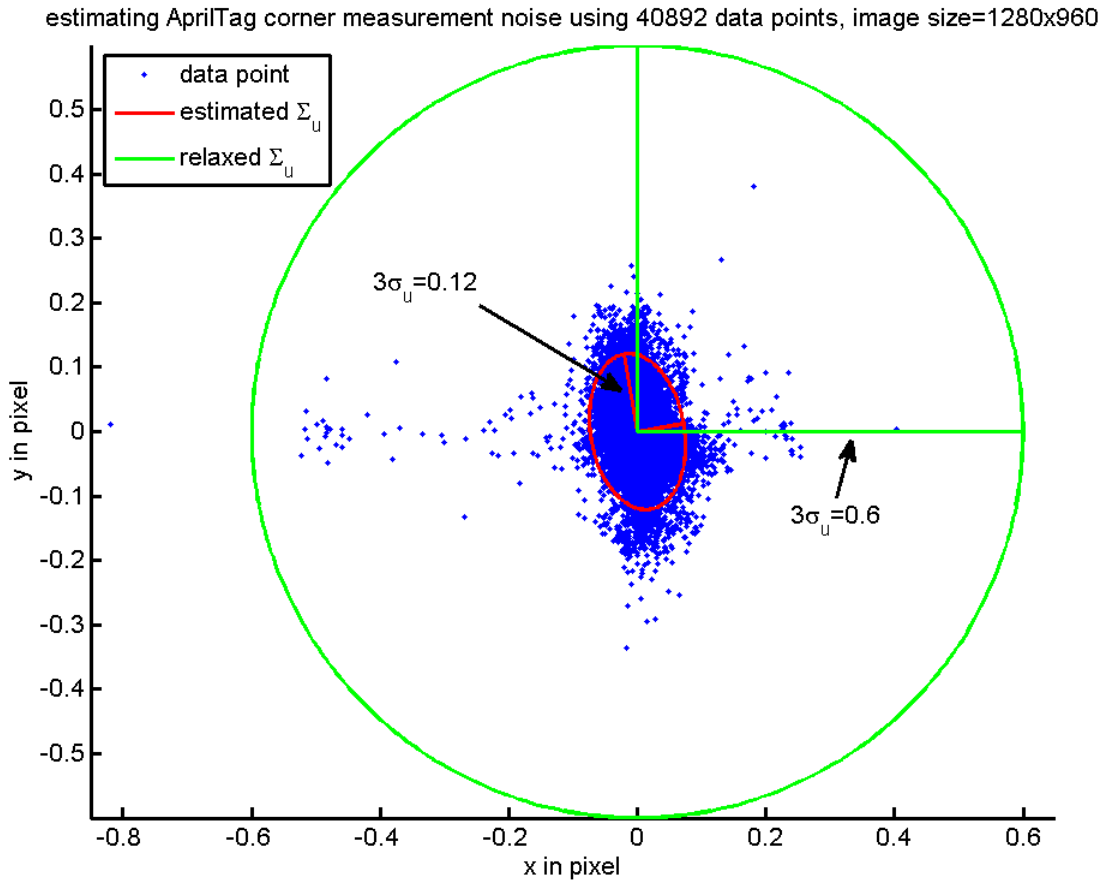
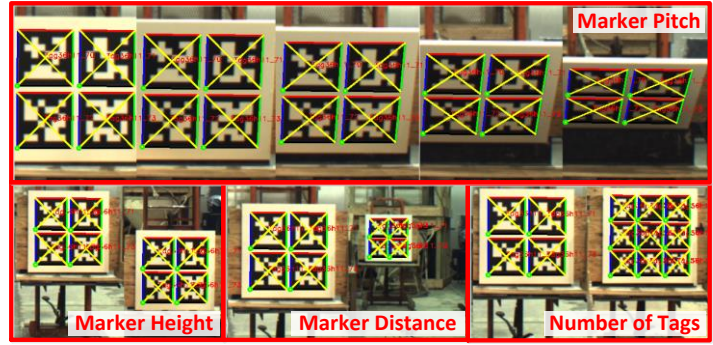


Figure 27. Image measurement noise estimation.



(a) Setup



(b) Different configurations

Figure 28. Mockup experiment setup.

Figure 29 shows the absolute depth errors comparing the ground truths with the results from camera marker network pose estimation, in the above mentioned different sets of prototype experiments, using the box quartile plot. Note that **all errors are less than 2.54 cm, even when observed from more than 10 meters away**. Further experiments showed that the system worked up to 15 meters.

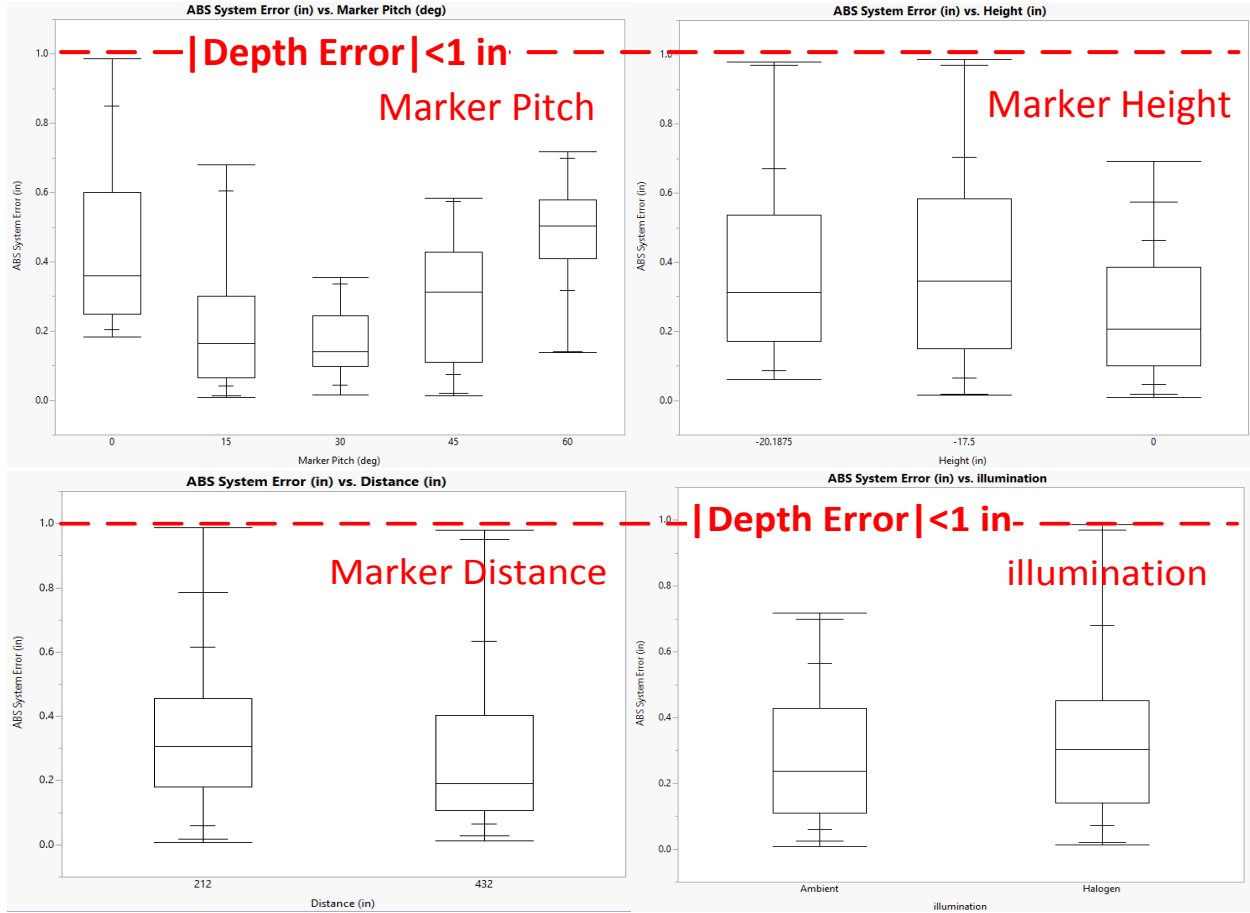


Figure 29. Error vs. configuration.

4.2.8 Field Demonstration of Articulated Machine Pose Estimation System

Another experiment was conducted to characterize the performance of the cable potentiometer prototype. The prototype was installed and tested on a Caterpillar 430E IT Backhoe Loader, as shown in Figure 30. The experiment involved calibrating the sensor system, placing the backhoe in random poses, using the sensors to estimate bucket tooth height, and comparing estimates with ground truth measurements obtained using a total station. The experiment's pass/fail criterion was set at 2.5 cm (1 in.) of absolute error. The operator was provided with a graphical user interface that shows the distance between the bucket teeth and the target profile. When the distance is less than 1 inch, the color switches from yellow to green.

A total of eight trials were conducted. For each trial, three components of tooth position (x , y , and z , where y corresponds to the zenith direction) were measured and compared with ground truth measurements, as shown in Figure 31. Of the twenty-four data points collected, only three points exceeded the pass/fail criterion of 2.5 cm (1 in.) of absolute error.



Figure 30. Calibration of cable potentiometer prototype installed on a backhoe.

Though the system's ability to estimate x -position only marginally met the pass criterion, the system's performance as a whole was deemed satisfactory, especially considering that accuracy along the zenith direction is more important in many excavation applications.

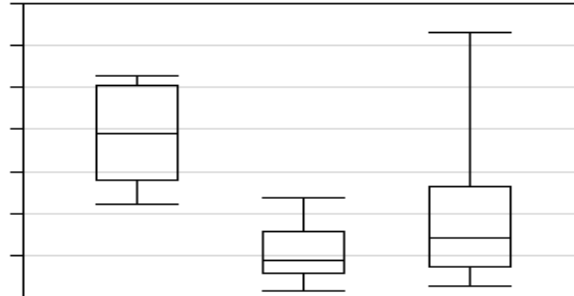


Figure 31. Cable potentiometer experimental results.

The cable potentiometer prototype was also tested on an active construction site, as shown in Figure 1 and Figure 32. The system was used to assist the operator in trenching operations. A computer screen was mounted in the operator's cabin, providing a display of bucket height relative to a desired trench grade specified by the job plans.



Figure 32. Cable potentiometer prototype installed on excavator.

Shown in Figure 33 is a short section of trench in which the operator conducted a side-by-side comparison of traditional grading versus grading guided by the sensor system. The resulting trench depth differences between the manual grade and the guided grade (by the prototype) were less than 1 in., which fulfils the need of many ROW excavation applications.

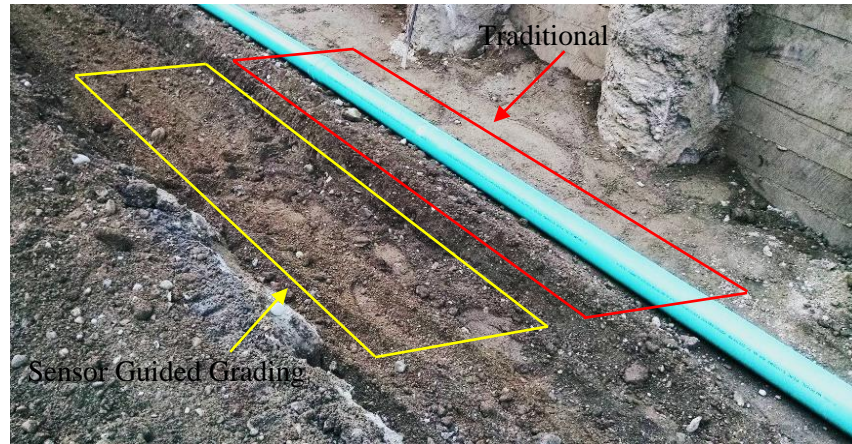


Figure 33. Comparison of sensor-guided grading versus traditional grading.

In summary, this stage of the project designed a vision-based pose estimation solution for articulated machines using a camera marker network. The uncertainty of the network pose estimation was analyzed through backward propagation of measurement covariance matrix. Based on this, an efficient approach of evaluating such a pose estimation system's uncertainty at any given state is introduced and applied to the basic single-camera single-marker system to find some important relationships between system states and corresponding system uncertainty, which is useful to guide more complex design. The conducted experiments and a working prototype proved the proposed solution's feasibility, robustness, and accuracy for real world ROW excavation applications.

Overall, the project solved two challenges in building a georeferenced augmented reality capability that allows an excavator operator to be visually aware of buried assets. The first one is to track the orientation and location of the excavator using a cost-effective camera marker network approach. Both simulation and field experimentation support the feasibility of the system which can achieve the accuracy of 1 in. within 40 m. In addition to tracking the relative distance between the bucket and the buried assets, this can also eliminate the need for grade checkers for excavations supporting water mains and storm sewer installations. The second challenge of Augmented Reality display is also addressed with accurate tracking of the excavator pose. The open issues of benchmark setup time lead to the future research needs of base marker network where a cluster of base markers can be deployed throughout the site and pre-calibrated/surveyed.

5 PLANS FOR IMPLEMENTATION

This IDEA project was conducted in collaboration with multiple industry stakeholders that included the DTE Energy Company (Michigan's largest electricity and gas provider), Miss Dig System (MDS) (Michigan's One-Call excavation safety company), Eagle Excavation (a specialty excavation contractor), and Walbridge (a large general contractor). Many of these collaborators are also associated with the Michigan Infrastructure & Transportation Association (MITA), which is a statewide construction trade association that consists of nearly 600 Michigan companies representing construction disciplines such as road and bridge, sewer and water, utility, railroad, excavation, and specialty construction throughout the state of Michigan. The industry collaborators have provided the research team with expert guidance on how the developed technology would impact their established member markets, and access to their excavators on construction sites for experimentation described in this report. The participation of the industry stakeholders has been guiding the knowledge transfer and commercialization process for the IDEA products explored in this project.

The products of this IDEA project are being developed further by a Michigan start-up company into a system called SmartDig (Figure 34). The SmartDig will assist excavator operators as follows:

1. The excavation solution will present operators with a visualization of excavation job plans, target grade profiles, and evolving grade profile in real-time. Persistent visualization of excavator job plans, target grade profiles, and evolving grade profile will be achieved through the use of AR technology explored in this project, thus rendering them in a spatially intuitive manner. This will allow operators to achieve target grades with high precision (within 2 in.) with fewer grade-checkers (75% to 80% less per job site) and improved grading productivity (5% to 8% increase per machine). It must be noted that the desired precision is a function of excavation type (e.g., water line, sewer) as well as applicable standards. It was found that 1 to 2 in. is a high standard for water line but a moderate standard for sewers. The operator confirmed the accuracy of SmartDig is in line with the requirement of most sewer installs except sanitary sewer, where the requirement is less than 1/8 in. (3mm) and this is currently achieved by shooting laser from laser level placed inside a manhole. However, SmartDig does not target sanitary sewer at the beginning. Additional feedback related to the graphical user interface as well as the display size which should be small enough so as not to obstruct the view sight of the operator.
2. The excavation solution will also present excavator operators with a persistent visualization of known utilities buried in their vicinity. Additionally, it will continuously monitor the proximity between the excavator's digging implement and these utilities, and will warn operators when they are about to strike these utilities. In the long run, the accidental utility strike warning features of the solution have the potential to save millions of dollars in liability and opportunity costs for excavation stakeholders.
3. The one-time setup to retrofit a machine with the developed technology components takes about 1 hour, and the time to set up a benchmark on a daily basis takes about 10 minutes. This time can be recovered

through the reduction of halting time of the excavator waiting for grade check results and rework if a target profile is not reached. In the future, if a network of bench/base markers can be installed throughout the site and pre-calibrated/surveyed, the daily setup of the benchmark can also be saved. The savings in operator and grade checkers' time is estimated at about 1hr/day on average, thus the savings of salary and fringe benefits of one operator and one grade checker in total is about \$100/day. A crew can be estimated to work for about 20 days/month and 9 months/year. The estimated annual saving is thus \$18,000 in total. Given the cost of SmartDig is estimated to about \$20,000, a contractor can recover their investment in about a year.

6 CONCLUSIONS

This IDEA project was conducted on the premise that ROW excavation safety will improve if equipment operators can persistently “see” where they are digging relative to expected utilities, and are objectively aware of the uncertainty associated with the utility locations. The exploratory project focused on creating and evaluating two key capabilities: (1) Persistent visualization of assets buried in an excavator's vicinity using a georeferenced augmented reality approach; and (2) Real time monitoring of an excavator's proximity to underground utilities using a graphical emulation approach. These capabilities were successfully created and were shown to be feasible in allowing an excavator operator to be visually aware of buried assets in a machine's vicinity, and providing a real time quantitative measure of a machine's distance to vicinal obstructions, significantly reducing the potential of buried utility strikes in ROW excavation. Accidents involving excavator hits to utilities are a long-standing and significant societal problem that leads to an unacceptable number of fatalities, injuries, property damage, and other costs each year (Sterling et al. 2009; Spurgin et al. 2009). Inadvertent utility strikes disrupt life and commerce, and pose physical danger to workers, bystanders, and the general public. The explored innovations thus have the potential to transform excavator operation from a skill-based to a knowledge-based process so future accidents are prevented.



Figure 34. SmartDig: (A) Camera cluster and stick marker; (B) Benchmark with pre-surveyed pose in the project reference frame; (C) System calibration; (D) Working prototype of automatic grade control; (E) Comparison to manual grade.

Although several variations of experimental prototypes were demonstrated with success, a robust system for large-scale deployment on actual construction sites has some additional technical challenges that need to be overcome. Future work is needed to identify the proper implementation of marker-based sensors in excavator pose estimation systems. For example, it is still unclear whether it is best to place cameras on an excavator in view of a job site, or on a job site in view of an excavator. One potential advantage of the former is the centralization of power, sensors, processors, and interface, while a potential advantage of the latter is the possibility of viewing multiple machines with the same camera. Also regarding implementation, experiments have indicated that it may be advantageous to employ a combination of marker-based sensors and traditional sensors, but the best combination and specific role of each sensor type must be further explored.

Another challenge needing to be addressed is camera-marker occlusion. Considering the busy nature of a construction site, it is not unreasonable to expect that a person, machine, or object could occasionally block the line of sight between a camera and marker. Thus, it may be necessary to evaluate the severity of the concern and address it, possibly through the incorporation of redundancies. Similarly, as an excavator moves, it may prove challenging to keep markers in the cameras' field of view. Again, one solution might exist in the use of multiple site-based cameras or multiple site-based markers for redundancy.

Lighting conditions may also pose a challenge, as low light levels can lead to marker detection loss. Thus, some form of active lighting may prove necessary for nighttime operation, and possibly even during dawn and dusk or in poorly lit indoor environments. Lastly, the harsh environment of the typical job site may present challenges for the sensor system's components. Thus, a fully functional system will require additional design efforts to ensure electrical and mechanical robustness.

7 REFERENCES

- Anspach, J.H., *The Case for a National As-Building Standard*, American Society of Civil Engineers, Reston, Va., 2011.
- AWWA, *Buried No Longer: Confronting America's Water Infrastructure Challenge*, American Water Works Association, Denver, Colo., 2012, pp. 3–4.
- Borthwick, J., P. Lawrence, and R. Hall, "Mining Haul Truck Localization Using Stereo Vision," *Proceedings of the Robotics and Applications Conference*, Vol. 664, 2009, p. 9.
- Bossler, J.D., J.B. Campbell, R.B. McMaster, and C. Rizos, *Manual of Geospatial Science and Technology*, 2nd ed., CRC Press, Boca Raton, Fla., 2010.
- Bowman, J. and P. Giovannoni, *CGA DIRT Analysis and Recommendations*, Vol. 10, Common Ground Alliance, 2013, pp. 2–4.

Brookshire, J., *Articulated Pose Estimation via Over-parametrization and Noise Projection*, MIT, Cambridge, Mass., 2014.

Bureau of Labor Statistics, *Census of Fatal Occupational Injuries (CFOI)—Current and Revised Data*, 2013 [Online]. Available: <http://www.bls.gov/iif/oshcfoi1.htm> [accessed Feb. 3, 2015].

CGA, Damage Prevention, 2015 [Online]. Available: <http://www.commongroundalliance.com/damage-prevention>.

Cheng, P. and B. Oelmann, “Joint-angle Measurement Using Accelerometers and Gyroscopes—A Survey,” *IEEE Transactions on Instrumentation and Measurement*, Vol. 59, No. 2, 2010, pp. 404–414.

Cho, Y. and M. Gai, “Projection-Recognition-Projection Method for Automatic Object Recognition and Registration for Dynamic Heavy Equipment Operations,” *Journal of Computing in Civil Engineering*, Vol. 28, No. 5, 2014, A4014002.

Common Ground Alliance, “Survey Finds That nearly Half of Homeowners Who Plan to Dig This Year Will Put Themselves and Others at Risk by Not Calling 811 Before Starting,” Mar. 3, 2015 [Online. Available: <http://commongroundalliance.com/media-reports/press-releases/survey-finds-nearly-half-homeowners-who-plan-dig-year-will-put> [accessed May 18, 2015].

Cui, Y. and S. Ge, “Autonomous Vehicle Positioning with GPS in Urban Canyon Environments,” *IEEE Transactions on Robotics and Automation*, Vol. 19, No. 1, 2003, pp. 15–25.

Duff, E., Tracking a Vehicle from a Rotating Platform with a Scanning Range Laser,” *Proceedings of the Australian Conference on Robotics and Automation*, Vol. 12, 2006.

Feng, C. and V.R. Kamat, “Plane Registration Leveraged by Global Constraints for Context-Aware AEC Applications,” *Computer-Aided Civil and Infrastructure Engineering*, Vol. 28, No. 5, 2013, pp. 325–343.

Feng, C., S. Dong, K.M. Lundeen, Y. Xiao, and V.R. Kamat, “Vision-Based Articulated Machine Pose Estimation for Excavation Monitoring and Guidance,” International Symposium on Automation and Robotics in Construction (ISARC), ISARC, Oulu, Finland, 2015.

Feng, C. and V.R. Kamat, “Plane Registration Leveraged by Global Constraints for Context-Aware AEC Applications,” *Computer-Aided Civil and Infrastructure Engineering*, Vol. 28, No. 5, 2012, pp. 325–343.

Feng, C., Y. Xiao, A. Willette, W. McGee, and V.R. Kamat, “Towards Autonomous Robotic In-Situ Assembly on Unstructured Construction Sites Using Monocular Vision,” *Proceedings of the 31th International Symposium on Automation and Robotics in Construction*, Sydney, Australia, 2014, pp. 163–170.

Hartley, R. and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, Cambridge, U.K., 2000.

Hel-Or, Y. and M. Werman, *Model Based Pose Estimation of Articulated and Constrained Objects*, ECCV, Springer, 1994, pp. 262–273.

Infrastructure Resources LLC, One Call and State Law Directory (USA), Excavation Safety Guide and Directory, Vol. 1, 2014, pp. 49–55.

Kashani, A., W. Owen, N. Himmelman, P. Lawrence, and R. Hall, "Laser Scanner-based End-effector Tracking and Joint Variable Extraction for Heavy Machinery," *The International Journal of Robotics Research*, Vol. 29, No. 10, 2010, pp. 1338–1352.

Lee, S., M.-S. Kang, D.-S., Shin, and C.-S. Han, "Estimation with Applications to Dynamic Status of an Excavator Without Renovation," *Gerontechnology*, Vol. 11, No. 2, 2012, p. 414.

Leica, Machine Control, Digging Solutions, 2015 [Online]. Available: http://www.leica-geosystems.com/en/Digging-Solutions_63797.htm.

Lin, L.-H., P. Lawrence, and R. Hall, "Robust Outdoor Stereo Vision SLAM for Heavy Machine Rotation Sensing. Machine Vision and Applications," Vol. 24, No. 1, 2013, pp. 205–226.

Memarzadeh, M., A. Heydarian, M. Golparvar-Fard, and J. Niebles, "Real-time and Automated Recognition and 2D Tracking of Construction Workers and Equipment from Site Video Streams," International Workshop on Computing in Civil Engineering, 2012.

Olson, E., "AprilTag: A Robust and Flexible Visual Fiducial System," *Proceedings of the 2011 IEEE International Conference on Robotics and Automation*, 2011, pp. 3400–3407.

Patel, A. and A. Chasey, "Integrating GPS and Laser Technology to Map Underground Utilities Installed Using Open Trench Method," 2010 Construction Research Congress, 2010, p. 628.

PHMSA, Pipeline Incidents by Cause, Pipeline and Hazardous Materials Safety Administration, U.S. Department of Transportation, Washington, D.C., 2015.

Rezazadeh Azar, E. and B. McCabe, "Part Based Model and Spatial—Temporal Reasoning to Recognize Hydraulic Excavators in Construction Images and Videos," *Automation in Construction*, Vol. 24, 2012, pp. 194–202.

Talmaki, S.A., *Real-Time Visualization for Prevention of Excavation Related Utility Strikes*, PhD Dissertation, Civil Engineering, University of Michigan, Ann Arbor, 2012.

Transportation Equity Act for the 21st Century, Vol. 112, U.S., 1998.

Trimble, Grade Control for Excavators, 2015 [Online]. Available: <http://construction.trimble.com/products/machine-control/grade-control-for-excavators>.

U.S. DOT PHMSA, *Pipeline Incident 20 Year Trends*, 2015 [Online]. Available: <http://www.phmsa.dot.gov/pipeline/library/datastatistics/pipelineincidenttrends> [accessed Feb. 4, 2015].

Yang, J., P. Vela, J. Teizer, and Z. Shi, "Vision-based Crane Tracking for Understanding Construction Activity," *Proceedings of the ASCE IWCCE*, 2011, pp. 258–265.