HUMAN-ROBOT INTEGRATION FOR POSE ESTIMATION AND SEMI-AUTONOMOUS NAVIGATION ON UNSTRUCTURED CONSTRUCTION SITES

*Chen Feng, Nicholas Fredricks, and Vineet R. Kamat
Civil and Environmental Engineering Department, University of Michigan
2350 Hayward Street
Ann Arbor, MI 48109
*(*Corresponding author: cforrest@umich.edu)*

HUMAN-ROBOT INTEGRATION FOR POSE ESTIMATION AND SEMI-AUTONOMOUS NAVIGATION ON UNSTRUCTURED CONSTRUCTION SITES

ABSTRACT

Compared to widespread successful deployment of robotic manipulators for repetitive and hazardous tasks in related industries such as manufacturing, the construction industry has achieved relatively limited benefits from robotics and soft automation. Unlike manufacturing, where robotic solutions benefit from the structured layout of the environment (e.g., factory assembly line), construction robots face unique challenges that arise from the rugged, dynamic, and unstructured environment of the work site, as well as the uncertainty and evolving sequence of occurring on-site events. This challenges any intended construction robots to not only replicate basic human motion, but also be capable of sensing and adapting to environmental changes, and making decisions based on the evolving state of the environment. Building upon recent advancements in robotic mapping, computer vision, and object recognition, the authors propose to introduce autonomous behavior at the basic task level for on-site construction robots to address these challenges in a flexible and extensible manner. This paper reports the outcome of the first phase of this research - a structured methodology for improved design and development of basic task automations - and focuses on algorithms developed for mobile robot navigation and relative pose estimation. The algorithms are implemented on a prototype mobile robotic platform, and evaluated in several experimental scenarios.

KEYWORDS

Semi-autonomous navigation, Pose estimation, Mobile robot, Construction basic task

INTRODUCTION

Robotics and automation in construction (RAC) is comprised of two major categories: hard and soft RAC (Balaguer, 2004). Just as “every construction chore has physical components and information components” (Everett & Slocum, 1994), hard RAC focuses mainly on construction tasks which contain a large portion of physical processing, such as robotics for brick laying, interior finishing, road paving, etc., while soft RAC concentrates mostly on construction tasks which typically require higher level information processing, such as document management, progress monitoring, safety monitoring, maintenance and inspection, and as-built Building Information Modeling (BIM).

However, even though hard RAC had been actively studied in the 1990s, RAC research has been shifting towards the soft RAC side since the last decade. From a previous research trend study (Son et al., 2010) about papers published in the proceedings of the International Symposium on Automation and Robotics in Construction (ISARC), a huge net decrease of hard RAC related papers from about 70% to 35% was observed. Part of the reason for this decrease is suspected be the following: unlike manufacturing, where robotic solutions benefit from the structured layout of the environment (e.g., factory assembly line), construction robots face unique challenges that arise from the rugged, dynamic, and unstructured environment of the work site, as well as the uncertainty and evolving sequence of occurring on-site events.

This trend and its reasons highlight the importance of incorporating more soft RAC techniques into the hard RAC side, which means more automatic information processing abilities should be developed for hard construction robots to increase their level of automation and thus make them easier to use (Balaguer, 2004). Meanwhile, building on recent advancements in robotic navigation and mapping, computer vision, and object recognition, researchers from robotics community have been making efforts towards autonomous robots that can perform certain simplified construction tasks, such as structure or brick assembly by quadrotors (Lindsey et al., 2012; Willmann, et al., 2012).

Based on the above analysis, and a brief review of related work, the authors propose a methodology for developing hard RAC techniques to allow more efficient implementation of autonomous behavior at the basic task level in on-site construction robots. Automatic pose estimation and autonomous navigation are identified as the core functionalities based on the analysis. Then, several research directions are described, summarizing the authors' previous and ongoing research efforts on the two core functionalities led by this methodology. Conclusions are then drawn in the final section of the paper.

RELATED WORK

Many RAC methodologies have been proposed as guidelines to identify construction tasks and develop robotics and automation solutions for them. Everett described a hierarchical taxonomy of construction field operations, in which two important levels of construction operation, activity and basic task, are proposed (Everett, 1991). While many hard RAC research had focused on activity level automation, i.e. whose output will "results in a recognizable, completed unit of work with spatial limits and/or dimensions" (Everett & Slocum, 1994), Everett (1991) proposed to conduct RAC research on the level of the basic task – fundamental elements that build up construction activities, since technology advancement on this level could be applied to many different construction activities, as opposed to automation on activity level. This paper follows the same idea and advances it in the next section by transposing the perspective of basic task level automation from construction worker to autonomous/semi-autonomous construction robots.

At the same time, as previously mentioned, a common reason for the difficulties in hard RAC is that the construction job-site environment is usually rugged and unstructured with uncertain events manifesting. Fully autonomous construction robots seem to lack both the required theoretical foundations and practical feasibility. Thus, semi-automation in construction enabled by human robot integration (HRI) is identified to be "preferential in the mobile and non-standardized construction environment" (Han, 2011). Previous works about either interior finishing robot (Kahane & Rosenfeld, 2004; Navon, 2000), where human operators need to manually transfer the robot between workstations, or infrastructure inspection and maintenance robot (Kim & Haas, 2000), where manual editing and correction of automatic crack sealing error is needed, had followed this principle. This paper is also guided by the same principle and develops a prototype of human robot interface for easier construction mobile robot operation that could be applied in a broad class of construction activities.

Apart from HRI as a midway solution, many researchers have realized that to increase the level of autonomy for construction robots, the mapping and navigation abilities of the robot are essential (Beliveau et al., 1996; Forsberg et al., 1997; Shohet & Rosenfeld, 1997). However, the accuracy of simultaneous localization and mapping (SLAM) algorithm is found to be insufficient at that time for many construction tasks which require direct manipulation of construction materials or tools (Shohet & Rosenfeld, 1997). Some researchers suggested removing the autonomous navigation functionality and transferring robots between workstations manually, then performing either a coarse-to-fine calibration (Kahane & Rosenfeld, 2004) or carrying out an additional vision-based real-time quality assurance step (Navon, 2000). While in this paper the authors propose to combine the aforementioned HRI prototype and the authors' previous research on general pose estimation (Feng & Kamat, 2012b), the accuracy requirement could be met while maintaining semi-autonomous navigation ability to make the machine easier to operate.

METHODOLOGY

In Everett's hierarchical taxonomy of construction field operations (Everett, 1991), the basic task level – including connect, cover, cut, dig, finish, inspect, measure, place, plan, position, spray and spread – is the one recommended for most easy introduction of construction automation. Since basic task is the fundamental element of construction field work, successful automation on one basic task could more easily benefit many different construction activities.

However, when RAC researchers actually try to automate these basic tasks, one issue they will encounter is likely to be the sub-problem overlap. For example, to automate the “connect” basic task, the first question for the designer to ask might be “how to identify the objects to be connected”. Thus object detection and recognition is a sub-problem for this basic task. On the other hand, to automate the “cut” basic task, the same sub-problem of object detection and recognition must be addressed since the robot needs to know what object needs to be cut. Similarly, the question “where and in what pose should the object be positioned” must be answered for the robot to automate both “position” and “place” basic tasks.

It is thus interesting to note that the basic tasks were summarized and abstracted from construction activities from the perspective of a human worker or manager. It is indeed natural, obvious and easy to assign to human workers commands made up from these basic tasks, whereas commands for construction robots require specification of additional detailed information in forms that machines understand.

Therefore, inspired by the modularization thinking in Everett’s methodology and the identification of overlapping sub-problems, the authors suggest that to efficiently automate basic tasks, their common sub-problems should be investigated and automated first. By further examining these sub-problems, one can realize that most of them are related to the information processing. Hence, the construction basic task automation methodology that the authors propose is as follows:

1) for each basic task, identify input and output information;
2) find each commonly needed type of information and define an atomic function which outputs that information;
3) prioritize all atomic functions and selectively automate them;
4) automate or semi-automate basic tasks which require information output by automated atomic functions.

This methodology is in line with the previous trend analysis stating that more automatic information processing abilities (the atomic functions) must be possessed by hard construction robots. Guided by this methodology, firstly the commonly needed information is analyzed (see Table 1). From Table 1 one can see that position and orientation information are the most commonly needed information. Moreover almost all autonomous mobile robots need this information to navigate themselves to their destination. Thus the authors choose to automate its corresponding atomic function: pose estimation and autonomous navigation.

Table 1 – Commonly needed information for each construction basic task

Basic Task	Object Identity	Position and/or Orientation	Area/Region/Shape/Boundary
Connect	√	√	
Cover	√		√ (Region to be covered)
Cut	√	√ (pose of cutting tool)	
Dig		√	√ (Region to be dug)
Finish			√ (Region to be finished)
Inspect	√	√	
Measure	√	√	
Place	√	√	
Plan		√	
Position	√	√	
Spray		√ (pose of spraying tool)	√ (Region to be sprayed)
Spread		√ (pose of spreading tool)	√ (Region to be spread)

IMPLEMENTATION AND RESULTS

In the above section, the two core functionalities, pose estimation and autonomous navigation have been identified to be critical for construction basic task automation. In this section, firstly the authors' previous work in pose estimation is summarized; secondly by combining this general pose estimation solution with the aforementioned well-recognized HRI principle, a prototype of semi-autonomous navigation framework is proposed.

General Pose Estimation

The authors have developed a general pose estimation solution previously (Feng & Kamat, 2012b). This camera based approach can more easily and robustly provide orientation and identity information compared to other sensors. By taking advantage of the fact that many construction scenes contain an abundant number of planar regions, the solution essentially suggests attaching planar markers on those regions of interest. Then by applying the proposed KEG tracker, the robot's relative orientation R_{camera}^{marker} and translation t_{camera}^{marker} w.r.t. the marker can be estimated in real-time (about 20 frames per second) with position error as low as about 3 mm. If the marker's global orientation R_{marker} and translation t_{marker} is known, then the robot's global pose can be calculated as $R_{camera} = R_{marker} R_{camera}^{marker}$ and $t_{camera} = R_{marker} t_{camera}^{marker} + t_{marker}$. As shown in Figure 1, the pyramids visualize the estimated relative poses R_{camera}^{marker} and t_{camera}^{marker} . This same idea could lead to an automation solution to manual drifting correction in ubiquitous context-aware computing (Akula, et al., 2011).

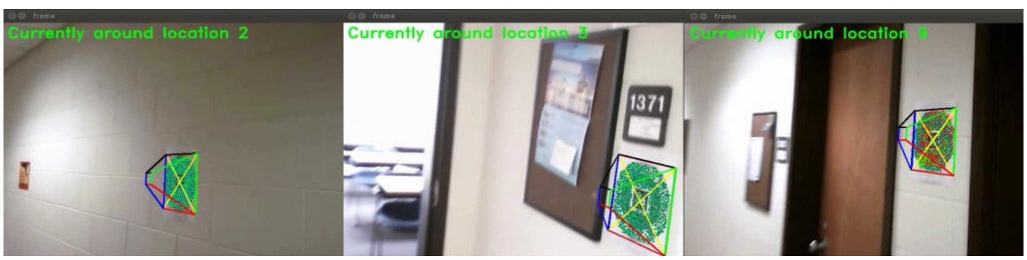

Figure 1 – Marker-based robot navigation

The authors then realized that by integrating identity information with the marker (Olson, 2011), certain level of object recognition ability is achieved. Moreover, with both identity and relative pose information, a more efficient methodology for assisting inspection and maintenance was proposed (Feng & Kamat, 2012a), treating markers as spatial indices so as to automatically retrieve associated information stored in the database (see Figure 2). This same methodology is currently being applied for enabling construction robots to autonomously identify and manipulate materials or tools to achieve basic tasks such as "connect" and "position" in the authors' Laboratory for Interactive Visualization in Engineering (LIVE) at the University of Michigan.

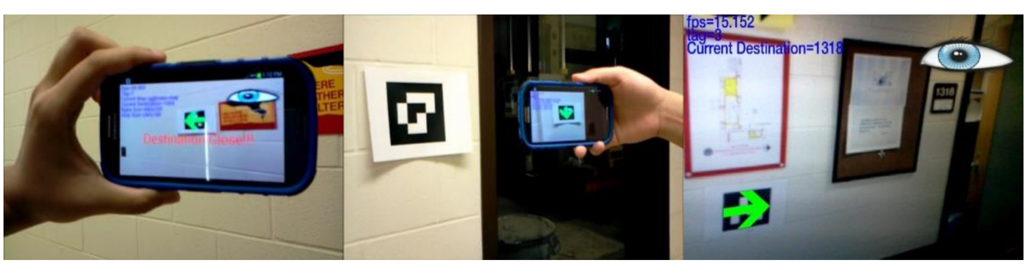

Figure 2 – Object identification and pose estimation for indoor inspection and maintenance

Human Robot Collaboration for Semi-autonomous Navigation

Different from the above scenarios where the markers are attached to static planar regions such as walls, floors or ceilings, the authors then extend this general pose estimation to circumstances where markers are to be attached to moving objects such as a human operator. This implies that the marker global pose R_{marker} and t_{marker} is changing over time. Inspired by the HRI principle, if a construction robot/machine could autonomously approach certain dynamically moving markers, this potentially provides a natural and easy-to-operate interface for a human worker to control the robot without extensive training. The first-stage prototype, “Follow-Me”, is described below.

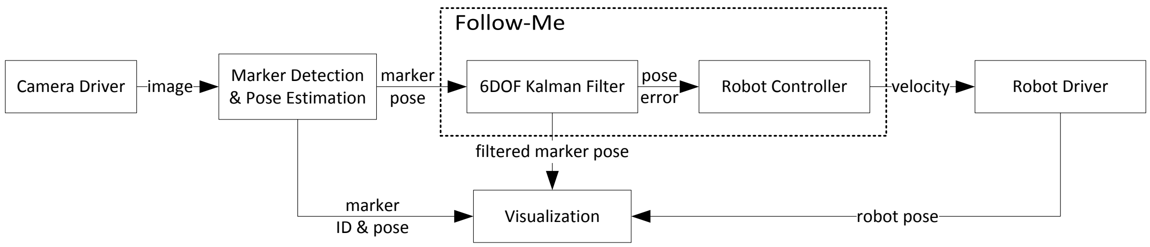

Figure 3 – Follow-Me system overview

The system is developed using the Robot Operating System (ROS) (Quigley, et al., 2009). As shown in Figure 3, the whole system contains 5 processes or nodes in ROS terminology. The camera driver node publishes images received from a webcam installed on the robot; the marker detection and pose estimation node subscribes those images and perform marker detection, which could either be running the AprilTag (Olson, 2011) or KEG (Feng & Kamat, 2012b) algorithm. If a marker with a certain ID is found, this node will publish the estimated pose information, which is then subscribed by both the Follow-Me node and a visualization node. The Follow-Me node is the “brain” of the system, containing two major components: a 6 degree of freedom (DOF) kalman filter for marker pose prediction in case that the marker is temporarily occluded or is not to be found, and a robot controller outputting wheel velocity command by taking the difference between the estimated/predicted marker pose and the reference pose (the ideal pose for the robot to be relative to the marker). Finally the robot driver node subscribes to the wheel velocity command, moves the robot accordingly and publishes the robot pose from its wheel encoder to the visualization node, which eventually takes the three poses and creates a 3D visualization of the scene.

The Follow-Me algorithm framework is shown in Figure 4. After initialization, the two major components will be triggered by different events inside an event loop. Whenever the Follow-Me node receives a new marker pose from the ROS infrastructure, the 6DOF kalman filter will be invoked to update the robot’s current knowledge about the pose of the marker. The other component, robot controller, will be activated for approximately every T_1 seconds. In order to handle various situations, the robot controller employs the behaviour based control principle (Arkin, 1998), containing three behaviours: predicting, exploring and following. If no new marker pose is received for any time between T_2 and T_3 seconds, the predicting behaviour is invoked and the robot will try to predict where the marker is going to be by the kalman filter. If no new marker pose is received for more than T_3 seconds, the robot will try to explore the nearby environment by spinning around itself, anticipating the marker. Other than these two situations, the robot will assume that the previously observed marker is still there and publish appropriate wheel velocity command through a proportional-integral-derivative (PID) control algorithm. In our experiments, $T_1 = 0.1$ second, $T_2 = 2$ seconds and $T_3 = 5$ seconds. Some preliminary results of the system are shown in Figure 5 (left image shows the working prototype; right image shows the trajectories of a human operator with a marker walking around a rectangular hallway followed by the prototype robot). Note that only a webcam on the testing robot is employed for sensing; other sensors such as a Kinect sensor are not used for this experiment.

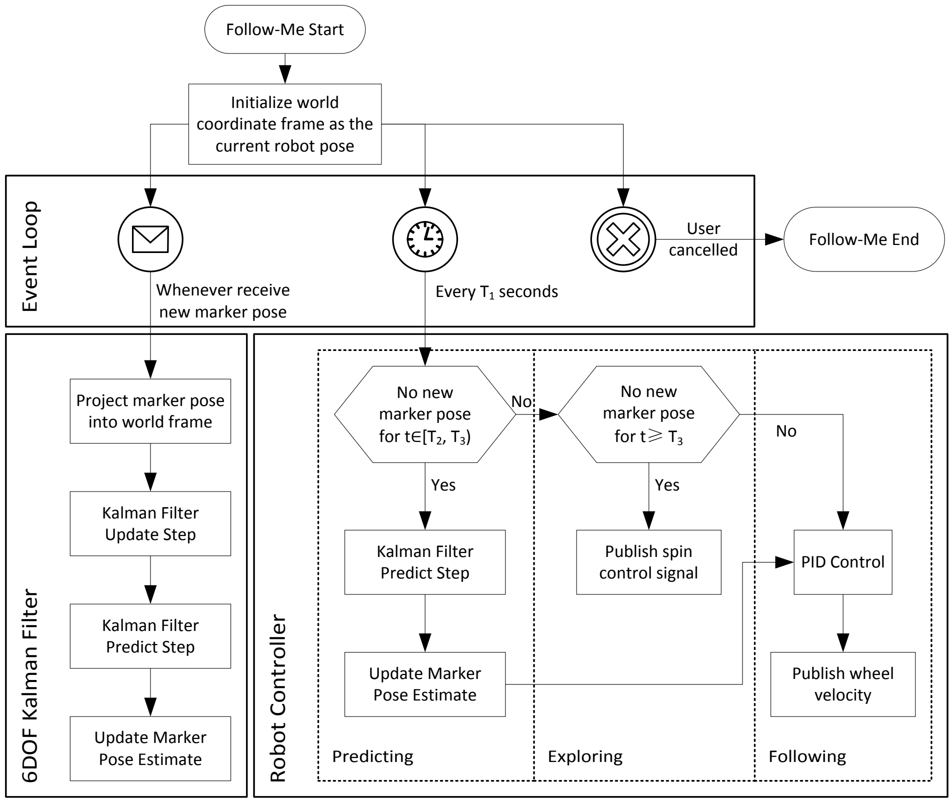

Figure 4 – Follow-Me algorithm framework

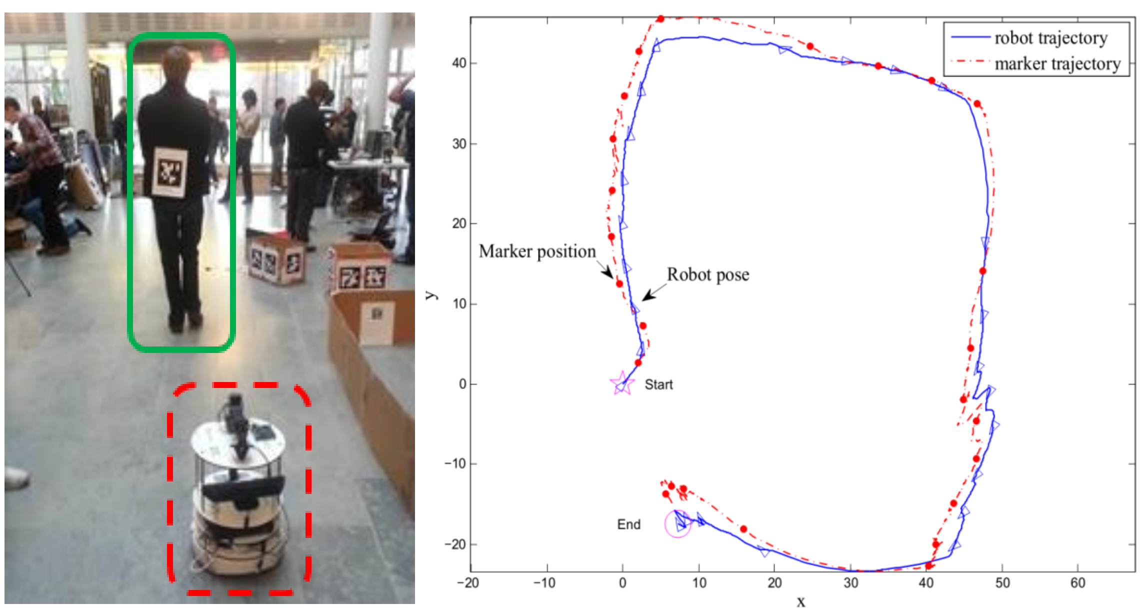

Figure 5 – Follow-Me system prototype working in a complicated environment (left: solid rectangle indicates the human operator with a marker; dash rectangle indicates the robot) and its trajectories

In summary, enabled by the above described two atomic functions, the following scenarios could be easier to achieve in a flexible and extensible manner. The construction robot will either autonomously navigate itself through pre-defined markers, or semi-autonomously steer itself by "Follow-Me" with a human pilot using a specific marker to reach its target workstation. During this procedure, lower SLAM accuracy of about 10 cm is enough since the goal for the robot is to roughly reach a zone around its final destination. After entering the zone, the robot can then detect markers inside and estimate its relative pose w.r.t. the marker with a much higher accuracy via the general pose estimation solution such as the KEG tracker. Using this highly accurate pose and additional identity information extracted from the marker, the robot is then ready to perform many basic tasks such as "position", "place", and "connect". Therefore, by following the HRI principle and the atomic-function identification methodology, the authors are able to implement hard RAC from the level of basic task.

CONCLUSIONS

This paper follows the idea of hard RAC at the basic task level, because this design method naturally leads to the function modularization and thus to flexible and extensible construction automation. Yet in order to automate construction basic tasks, the sub-problem overlapping issue is recognized by the authors. This issue should be addressed in order to facilitate a more effective design and development of basic task automations. Realizing the fact that most overlapping sub-problems are related to construction information processing, the authors propose a methodology to firstly collect and analyze commonly needed information for construction basic tasks, secondly define atomic functions producing those commonly needed information and then selectively automate some atomic functions. Guided by the methodology, two atomic functions, pose estimation and autonomous navigation, are chosen to be automated.

Through the authors' previous efforts on plane registration algorithm, the general pose estimation solution is then described briefly. After that a semi-autonomous navigation system prototype is developed, led by the HRI principle, which essentially enables a construction mobile robot to always drive itself autonomously towards its human pilot equipped with a specific marker. This provides a natural and easy-to-operate interface for human worker to control the robot without heavy training. In addition, by combining this with the highly accurate general pose estimation solution, the robot is then well-suited to perform many basic tasks in a coarse-to-fine manner. It is worth noting that for rugged construction environments, with vision based method providing vital information such as orientation and identity more easily and robustly, better autonomy can be achieved by combining different sensors, which could overcome the limitation of pure camera based solution such as occlusion and lighting requirement.

In the future, the authors plan to first improve the "Follow-Me" system's robustness by adding more behaviours such as collision avoidance using laser or 3D image sensor, local path planning, better marker prediction through particle filter etc. Secondly, the authors also plan to apply the same methodology for enabling construction robot to autonomously identify and manipulate materials or tools so as to automate basic tasks such as "connect" and "position". Thirdly, as an effort of moving from the HRI midway solution towards a higher level of autonomy in construction robot navigation and mapping, the authors plan to introduce more prior knowledge from construction domain to robotics SLAM problem to specifically develop SLAM algorithm targeted for construction scenarios.

REFERENCES

Akula, M., Dong, S., Kamat, V., Ojeda, L., Borrell, A., & Borenstein, J. (2011). Integration of infrastructure based positioning systems and inertial navigation for ubiquitous context-aware engineering applications. *Advanced Engineering Informatics, 25*(4), 640--655.

Arkin, R. C. (1998). *behaviour based control* . MIT Press.

Balaguer, C. (2004). Nowadays trends in robotics and automation in construction industry: Transition from hard to soft robotics. *In 21st International Symposium on Automation and Robotics in Construction (ISARC).* Jeju. Korea.

Beliveau, Y., Fithian, J., & Deisenroth, M. (1996). Autonomous vehicle navigation with real-time 3D laser based positioning for construction. *Automation in construction, 5*(4), 261--272.

Everett, J. (1991). *Construction automation--basic task selection and development of the CRANIUM.* Massachusetts Institute of Technology.

Everett, J., & Slocum, A. (1994). Automation and robotics opportunities: construction versus manufacturing. *Journal of construction engineering and management, 120*(2), 443--452.

Feng, C., & Kamat, V. R. (2012a). Augmented Reality Markers as Spatial Indices for Indoor Mobile AECFM Applications. *Proceedings of the 2012 Conference on Construction Applications of Virtual Reality (CONVR)*, (pp. 235-242). National Taiwan University, Taipei, Taiwan.

Feng, C., & Kamat, V. R. (2012b). Plane Registration Leveraged by Global Constraints for Context-Aware AEC Applications. *Computer-Aided Civil and Infrastructure Engineering*.

Forsberg, J., Aarenstrup, R., & Wernersson, A. (1997). A construction Robot for Autonomous Plastering of Walls and Ceilings. *Proceedings of the 14th ISARC*. Pittsburgh, Pennsylvania, USA.

Han, C.-s. (2011). "Human-Robot Cooperation Technology" An Ideal Midway Solution Heading Toward The Future Of Robotics And Automation In Construction. *Proceedings of the 28th ISARC*. Seoul, Korea.

Kahane, B., & Rosenfeld, Y. (2004). Balancing Human and Robot Integration in Building Tasks. *Computer Aided Civil and Infrastructure Engineering, 19*(6), 393--410.

Kahane, B., & Rosenfeld, Y. (2004). Real-time “Sense-and-Act” operation for construction robots. *Automation in construction, 13*(6), 751--764.

Kim, Y., & Haas, C. (2000). A model for automation of infrastructure maintenance using representational forms. *Automation in Construction, 10*(1), 57--68.

Lindsey, Q., Mellinger, D., & Kumar, V. (2012). Construction with quadrotor teams. *Autonomous Robots*, 1--14.

Navon, R. (2000). Process and quality control with a video camera, for a floor-tilling robot. *Automation in construction, 10*(1), 113--125.

Olson, E. (2011). AprilTag: A robust and flexible visual fiducial system. *Proceedings of the 2011 IEEE International Conference on Robotics and Automation (ICRA)*, (pp. 3400-3407).

Quigley, M., Gerkey, B., Conley, K., Faust, J., Foote, T., Leibs, J., . . . Ng, A. (2009). ROS: an open-source Robot Operating System. *ICRA workshop on open source software, 3*(3.2).

Shohet, I., & Rosenfeld, Y. (1997). Robotic mapping of building interior—precision analysis. *Automation in construction*, *7*(1), 1--12.

Son, H., Kim, C., Kim, H., Han, S., & Kim, M. (2010). Trend analysis of research and development on automation and robotics technology in the construction industry. *KSCE Journal of Civil Engineering, 14*(2), 131--139.

Willmann, J., Augugliaro, F., Cadalbert, T., D'Andrea, R., Gramazio, F., & Kohler, M. (2012). Aerial Robotic Construction Towards a New Field of Architectural Research. *International Journal of Architectural Computing, 10*(3), 439--460.