Conceptual Framework of Training Simulator for Heavy Construction Equipment
Integrating Sensory Data, Actual Spatial Model, and Multi-Agent System

Amin Hammad¹, Faridaddin Vahdatikhaki², Khaled El-Amamri³

1) Professor, Concordia Institute for Information Systems Engineering, Concordia University, Montreal, Quebec, Canada. Email: hammad@ciise.concordia.ca
2) Research Associate, Department of Building, Civil and Environmental Engineering, Concordia University, Montreal, Quebec, Canada. Email: vahdatifarid@yahoo.com
3) Graduate Research Assistant, Department of Building, Civil and Environmental Engineering, Concordia University, Montreal, Quebec, Canada. Email: elamm_k@encs.concordia.ca

Abstract:
The construction industry claims one of the highest occupational fatality rates. A great portion of onsite accidents is attributed to the unsafe and risky operations of heavy construction equipment, such as excavators. Such risky operations stem in many cases from inadequate training of heavy equipment operators. Given the high cost of training programs, where operators are trained on actual equipment, Virtual Reality (VR)-based training simulators are gaining popularity both in the industry and in academia. However, current simulators are limited in the available training scenarios and in capturing the dynamic interactive environment where several operators need to work in tight areas while ensuring smooth and safe operations. In the available simulators, training scenarios are developed manually and therefore there are usually a limited number of scenarios from which the trainees can choose. While VR environment is investigated for application in construction training simulator, advanced immersive technologies that can provide a more realistic training experience are not fully considered. Additionally, while the trainees are able to control, navigate and operate the VR equipment, the realistic behavior of surrounding pieces of equipment, with which the operator must interact, is scarcely simulated. Last but not least, the existing simulators are mainly designed to accommodate a single user at a time. Given that communication is an indispensable ingredient for effective and safe equipment handling, training simulators need to address multi-user environment where several trainees can simultaneously work, or collaborate, on a training scenario. In this paper, a novel conceptual framework is proposed to develop the training scenarios from actual construction sites by integrating sensory data captured from the site and an actual spatial model (including the Building Information Model and the terrain model). Additionally, it is proposed to develop a Multi-Agent System inside the training environment to simulate the behavior of multiple pieces of equipment surrounding the virtual equipment controlled by the trainee. These components are bound together in a multi-user immersive environment. A case study is conducted to demonstrate the feasibility of the proposed method. It is demonstrated that the proposed framework is providing a promising alternative to the current training simulators where the operators can be economically subject to a multitude of realistic training scenarios to develop both their operational and communication skills required for safe and productive operations.

Keywords: Construction Equipment Training, Immersive Virtual Reality, Simulators, Building Information Modeling, Multi-agent Systems.

1. INTRODUCTION
Construction projects, especially excavation projects, are notorious of being one of the highest sources of occupational fatalities and injuries in Canada and other counties. In earthwork operations, where heavy machines are used, various safety and risk issues could threaten the workers on site and delay the completion of the project. Additionally, the construction working environment is heavily susceptible to unforeseen changes and circumstances that could impact the safety conditions of the site. For instance, according to Safe Work Manitoba (2014), in the period 2000-2013, the acute-hazard exposure fatalities of the construction sector was 18% of all industries in Manitoba. Furthermore, the number of injuries in the construction sector increased between 2000 and 2013 and reached 14% of all sectors in 2013. It can be noticed from the same report that the accidents were not limited to novice workers, but in many cases involved skilled workers. Therefore, thorough training of the operators of heavy earthmoving equipment (e.g., excavators) is of high importance. In addition to training on actual equipment, new information technologies are providing computerized training simulators using Virtual Reality (VR) environments, which are economical and safer than training on actual equipment. However, current heavy construction equipment simulators use a non-immersive Virtual Environment (VE) to build the basic skills of equipment operators only for one user at a time using predefined and customizable scenarios. Additionally, little attention is paid to capturing and simulating the behavior of the surrounding equipment and workers so as to provide a more life-like training experience. Accordingly, there is a need for another level of training that aims to familiarize operators with actual site conditions of projects where they need to interact with other pieces
of equipment and react to complex and unpredictable situations.

This research aims to investigate and develop the next generation simulator for heavy equipment training (called Future Heavy Construction Equipment Simulator-FHCES) that can be used to train multiple trainees for equipment handling, safety, and communication in an immersive VR environment using serious gaming scenarios that reflect actual site conditions (e.g. congestion level). The safety training will consider the interaction of the trainee with other pieces of equipment controlled by software agents and his/her ability to react under complex and unpredictable situations. Focusing on excavation operations, the proposed project aims to improve the training experience using a Multi-agent System (MAS) and serious gaming.

2. LITERATURE REVIEW

2.1 Training Simulation and Serious Gaming in Construction Industry

Simulation is an active area of research for training heavy equipment operators in a safe VE. For example, Proctor et al. (2012) presented a Hierarchical Task Analysis (HTA) for excavator digging a trench and loading a truck. The HTA includes the plans showing the conditions under which each of the sub-tasks are triggered and whether the subtasks are performed sequentially or concurrently.

Although a simulated environment is not a full substitute for actual experience using real equipment in a work setting, simulation technology offers the following advantages (Canadian Apprenticeship Forum 2013): (1) Simulators are sometimes less expensive than actual equipment and consume less energy; (2) Simulators are safer, which is especially important for novice trainees; (3) Simulators detect and correct errors before they become habits; (4) Dexterity skills are acquired at least as efficiently using simulator technology and may be acquired more quickly; and (5) The level of complexity can be altered using a simulator so that tasks become progressively more challenging. The Operating Engineers Training Institute of Ontario (OETIO) was the first training center in the world to use a simulator to teach crane operation (Canadian Apprenticeship Forum 2013).

On the other hand, the recent progress in video game engines encouraged researchers to build serious games for training purpose (Juang et al. 2011). A serious game is a game designed for critical purposes such as defense, education, emergency management, city planning, and engineering (Sherif and Mekkawi 2009, Wang et al. 2011). In addition, serious games are usually simulations of real-world phenomena including events or processes designed for solving real problems. In addition to being entertaining, serious games are geared to train and educate users in an interactive simulation environment. In recent years, serious games became popular in Architecture, Engineering and Construction (AEC) research. Chavada et al. (2012) presented an approach for integrating workspace management within the planning process using a game engine. Miller et al. (2012) discussed the development of a virtual training environment to provide construction managers and trainees with a means to experience health and safety issues as they occur in a real construction site. Ruppel et al. (2011) discussed the design of a serious game based on Building Information Modeling (BIM) for fire safety evacuation simulation. Serious games have been proposed as an effective tool for safety training in the AEC domain. Lin et al. (2005) developed an educational game where the players assume the role of construction safety inspectors and should identify safety violations while moving in the construction site. Zhao et al. (2008) developed a VR environment for electrical safety awareness and training using the Torque game engine.

Beside interaction, visualization is an important part of any serious game. There are different visualization modes such as Virtual Reality (VR) and Mixed Reality (MR). These two modes are used to allow system users to visually interact with virtual objects. In the MR mode, the user can interact with the virtual objects augmenting the scene using a graphical interface. For example, Halbach and Halme (2012) presented a concept for an MR-based job planning interface that human operators could use to specify plans for automated earthmoving. The augmentation is a result of combining an updated 3D workspace model with the worksite images. This interface provides only geometrical and location information and does not consider real-time interaction. Moreover, Hammad et al. (2009) investigated the effectiveness of distributed MR for visualizing collaborative construction tasks in the context of the training of crane operators. However, high-end, military-grade Head-Mounted Displays (HMDs) were used in this work, which are not practical for the purpose of construction equipment simulator. Fang and Teizer (2014) presented a framework for using actual site data for the development of training VR environment. Nevertheless, this framework does not account for conflicts between the trainee-controlled equipment and the interacting surrounding equipment. Vasenev (2015) developed a VR environment for the visualization of roller operations using actual data from the site that can be further used to analyze alternative routes for the roller.

Based on the above review, there are several research works related to the usage of simulators for heavy equipment operation training and serious games applications in training. However, in the current state of the practice, the training scenarios are manually designed and built based on hypothetical or semi-hypothetical settings. This arrangement not only renders the development of training scenarios labor-intensive and expensive
but also results in limited scenarios that cannot fully represent the dynamism, intricacies and uncertainties intrinsic in construction sites. Also, the current construction equipment simulators use a single-user and non-immersive VE based on predefined and customizable scenarios. Therefore, more realistic training simulators are required to address the needs of the trainees to be exposed to the actual site conditions where the behavior of peripheral equipment/workers and all the uncertainties and intricacies inherent in earthwork sites are captured.

2.2 Multi-agent Systems

The main ingredients of MASs are agents capable of interacting with each other in an intelligent manner. An agent is defined as an entity situated in an environment with the capability to form a perception of the environment and act upon it, which in the context of Artificial Intelligence is materialized through the application of sensors and actuators (Russell and Norvig 2003). Intelligent agents can make decisions in tasks that require high computational efforts or complex coordination in a pure simulation environment or in a real-time system (e.g. Tambe 1997, Yen et al. 2006, Fan and Yen 2007). Although, the application of MASs for the simulation of construction operations has been investigated (Sawhney et al. 2003, Rojas and Mukherjee 2006, Van Tol and AbouRizk 2006, Marzouk and Ali 2013), the application of MASs as a decision-making and planning tool is very limited. Ren and Anumba (2002) developed a MAS framework with learning capabilities to facilitate the construction claim negotiations. In this framework, three types of agents are used to represent contractors, engineers and clients as the main participants of a claim negotiation. A negotiation strategy is proposed that calculates an agent’s likelihood of risk acceptability based on the amount of the utility the agent achieves in different scenarios. Kim and Russel (2003a and 2003b) proposed a framework for an intelligent earthwork system based on the application of a MAS architecture. In this framework, three main subsystems, namely task planning, task execution and human control, have been suggested to provide a complete coverage for an earthwork project task assignment and execution. The proposed framework, while providing a good structure for the distribution of responsibilities between a multi-layer agents hierarchy, assumes an un-deviated execution of the schedule and does not cover functionalities to monitor the actual progress of the operations with regard to the generated schedule in order to provide proactive corrective measures for keeping the project on schedule. Additionally, safety issues of construction sites are not addressed by the proposed MAS. In another effort, an agent-based system was proposed for the communication in construction projects by Lee and Bernold (2008). In this system, a set of agents are used to collect weather data from the Internet and on-site instrumentations, forecast hazardous weather conditions and warn crane operators.

3. PROPOSED METHOD

The proposed method is mainly established on the idea of generating the immersive training VR environment through the integration of a MAS with tracking data captured from the site, and the available spatial models, such as BIM and Digital Terrain Model (DTM). Figure 1 represents the main framework of the proposed method. At the high level of abstraction, the framework consists of five main steps, namely (1) Data Collection, (2) Data Preparation, (3) Training Scenario Generation, (4) User Interaction, (5) Feedback on Safety and Productivity Performance.

3.1 Data Collection

In the data collection step, a set of tracking technologies (such as GPS, Ultra Wideband (UWB), Radio Frequency Identification (RFID), etc.), laser scanners, cameras, and weather stations can be used to collect data about the activities and conditions on the site. These technologies are used to track equipment, workers and materials on the site and the changes in the weather conditions. Also, it is important to track the changes in the topography of the site using aerial and ground laser scanning. This information has a twofold application: (1) it can represent how the site conditions change with respect to ongoing earthmoving operations; (2) it can be used to enhance the a priori physical modeling used to represent the causal interactions between actions and their resultant reactions. It is noteworthy that the increasing popularity of applying Automated Machine Control and Guidance (AMC/G) technologies in construction sites can greatly streamline the process of data collection. This is due to the fact that the tracking data coming from high-accuracy Real-Time Location Systems (RTLS) used in AMC/G can be readily used for generating a realistic training VR environment.

3.2 Data Preparation

In the data preparation step, the data from the previous step are used to generate four main pieces of information, namely (1) the updated 3D model of the site (e.g., BIM), (2) the updated DTM, (3) the movements of workers, equipment, and material, and (4) the movement pattern of equipment. The 3D model of the site is built using the collected spatial data. This spatial model includes one or a combination of BIM, Road Information Modeling (RIM), Bridge Information Modeling (BrIM), etc., as well as other features such as underground pipes and cables that could be obtained from the GIS models of the site. The assumption of this method is that these spatial
elements can be integrated in a single model. The authors have previously discussed about generating such spatial models (Hammad et al. 2013). This integrated model can be updated using the data gathered from the previous step. The updated model would represent “as-is” status of the site at different points of time.

It should be highlighted that the difference between the tracking data and movement patterns is that the latter represents the underlying rules that govern the paths and motions of the different pieces of equipment. As such, the tracking information of equipment can be used to develop the higher-level behavior model for the equipment that can be used to develop realistic agents capable of capturing the operational logic of different equipment in different operations. For instance, the repetitive trips between the dumping and excavation areas and the duration of stay in each area can be used as the basis to extract the behavioral pattern of a truck in a typical earthmoving operation. The authors have previously demonstrated an example of such behavioral model for excavators (Vahdatikhaki and Hammad 2015b). These agents are parts of the MAS that handles the coordination of the individual agents for smooth and realistic interaction.

![Figure 1. The overview of the framework for the proposed method](image)

This data collection and integration method for building a georeferenced VE from several sources including, DTM, GIS, BIM and CityGML result in replicating the real world environment, which allows farther sensory data integration later on. In addition to the surrounding structures and operating equipment, the trainee is exposed to the hidden infrastructure (e.g., sewage and water pipes, underground wiring, etc.). These data can be visualized as an additional layer on his view based on his position to improve safety and task efficiency.

### 3.3 Generation of Training Scenarios

The training scenarios are generated combining the data from the previous step in a game engine environment. In creating the scenarios, the training specialist considers several factors: (1) The specialist needs to find an appropriate site from the library of various construction sites from where the adequate data are collected. In doing so, the specialist considers the geographic specifications of where the trainees are going to work and the nature of the work they are going to perform; (2) The specialist needs to determine the pieces of equipment that are going to be handled by the trainees. Depending on the number of trainees the scenario can accommodate, the specialist marks the equipment that will be controlled by the trainees; (3) At the next step, the specialist needs to review the captured data from the selected site and identify the portion of the activities that can be used to generate a training scenario (i.e., period within which the involved pieces of equipment were active). While it is important to ensure that the scenario contains activities of other equipment and workers on the site, this should not be misconstrued with eliminating the near-miss incidents. On the contrary, such safety-sensitive incidents provide a very effective and realistic exposure for the trainees to hone their skills in avoiding everyday risks on the site; and (4) The specialist should also consider that the level of difficulty of the scenario should be commensurate with the skill level of the targeted trainees. It should be also highlighted that the specialist can decide to exclude a certain number of equipment and workers from the scenario to adjust the congestion level and thus the level of difficulty. For instance, targeting a novice trainee, the specialist may remove a nearby excavation team (consisting of excavators and trucks) to allow the trainee to focus on improving his/her equipment handling skills. Additionally, remote equipment and workers, which are not going to impact the operation of the selected equipment, can be excluded to reduce the computational intensity of the training scenarios.

### 3.4 User Interaction

Once a scenario is built, trainees can start the training. The surrounding equipment and workers can be categorized into two groups. If the surrounding equipment/workers are not directly interacting with the selected equipment for the trainee(s), the actual data from the site (i.e., the paths) are used to represent their movements. On the other hand, if the surrounding equipment/workers are in direct interaction with the selected equipment, their movements and decisions are generated by the agents of the MAS. Concerning the level of detail of the
simulated or replayed workers, it is sufficient to represent their movements only as walking or static figures. In other words, capturing the postures and motions of workers only offer a limited advantage at the cost of additional requirements for complex data collection/preparation and computation power.

Conventionally, construction equipment simulators are equipped with 3-degree-of-freedom (3DOF) motion platform with surround audiovisual systems (CM Labs 2015). These platforms are able to provide a realistic exposure to the motion physics originating from the interaction of the equipment with the VR model. These platforms together with the surround audiovisual systems and realistic control units, including joysticks, create an immersive VR environment. While the existing platforms create an acceptable VR experience, the realism of the experience can be further enhanced using an immersive VR HMD (e.g. Oculus Rift (2015)). These displays offer a truly immersive environment using head tracking. This will allow the trainees to navigate in the VR environment by moving his/her head. The tracking range of the HMD’s 3 DOF movements (i.e., 3-axis rotational tracking) forms a pyramid-shaped Field of View (FOV). This volumetric FOV defines the range of tracking for the head tracking device. The HMD-attached gesture tracking device allows to track the movement of the operator hands and fingers and to virtually replicate them in the VE to assist him/her to locate the physical joysticks. Figure 2(a) shows the hardware components and Figure 2(b) shows the head and gesture tracking volumes.

As mentioned in Section 1, multiple users can interact in a single scenario in the proposed method. The multi-user compatible setting of the proposed method allows trainees not only to develop their dexterity and equipment handling skills but also their communication skills. According to Figure 2(c), which shows the trainees and the schematic representation of the scenes as seen through the HMD, multiple trainees can work in a same team (e.g., one trainee operates the truck and another operates the excavator) or in different teams (e.g., each trainee operates an excavator). When working in a team, trainees should be aware that their sluggishness and mistakes would impact the performance of the team. Also, they should be sensitized to safety aspects of their decisions, trying to avoid collisions and dangerous encounters with the trainee-controlled, MAS-controlled, and tracking-data-propelled equipment. If the trainees are working in different teams, they should primarily pay attention to avoiding conflicts with the operations of the equipment of the other team.

Another feature of the proposed method is the provision of adjustable level of hints and guidance for the trainees in form of visual instructions about how to steer the equipment, what task to do, how much soil to remove, and information about the surroundings. The Heads-Up Display (HUD) and the interactive GUI allow the operator to view warnings in the form of messages combined with moving arrows (e.g., dump truck approaching the work space from the right side with a speed of 10 km/hr). In addition, The GUI provides a real-time productivity report based on the tasks achieved and the excavated soil. This is similar to the visual guidance provided to the operators of AMC/G-enabled equipment. The advantages of this feature include: (1) providing the novice worker with step-by-step and easy-to-follow instructions, (2) providing safety guidance using the Dynamic Equipment Workspaces (DEWs) (Vahdatikhaki and Hammad 2015a) and Look-ahead Equipment Workspaces (LAEWs) (Vahdatikhaki and Hammad 2015b), and (3) enabling expert operators to be trained for AMC/G-enabled equipment and to hone their skills in using the provided guidance in an effective and active manner.
3.6 Feedback on Safety and Productivity Performances

Once the training session completes, the trainees are evaluated based on the equipment control and productivity (i.e., dexterity and smooth motions) and safety performance. The control and productivity performance is measured considering (1) the average cycle time of the operation, (2) the waiting time and length in the queue, (3) the length and the smoothness of the paths generated by the trainee, (4) the ability to control multiple DOFs simultaneously, and (5) the ability to coordinate with other trainees or agents. The safety performance, on the other hand, is evaluated based on (1) number of collisions, (2) near miss events calculated based on collisions between DEWs of equipment, and (3) ability to avoid hazardous encounters as indicated by LAEWs.

The proposed VR environment records the training session and marks parts of the performance where the trainees did not perform satisfactorily. The trainees can review their performances and observe their mistakes at the end of each training session.

4. IMPLEMENTATIONS AND CASE STUDY

As shown in Figure 3, a prototype training immersive VR environment for earthmoving operations is developed in Unity (2015) where two excavators and four trucks are working together as a team. This scenario is designed for only one trainee and all other pieces of equipment in the scenario are represented by agents in a MAS. At the present stage of the prototype, since the equipment tracking data from the site is missing, all surrounding pieces of equipment are represented by agents.

As shown in Figure 3(a), the case study was inspired by the Turcot interchange project (Transports Quebec 2014), which is a large interchange reconstruction project in the city of Montreal. The DTM of the area of the case study was converted to a TIN format using ArcGIS (2015). Then, it is exported to FBX file format and imported into Unity scene. It is then converted to the terrain type used by Unity (Height Map) to allow terrain deformation during excavation. The height Map terrain model is textured with an aerial photo. The 3D model of the existing highway bridges were developed in Bentley’s MicroStation (2015) and then imported into Unity in FBX format.

Figure 3(b) shows the equipment used in the case study and the site layout. The trainee is using two Saitek’s Cyborg Evo Wireless joysticks (2015) to control the 6 DOFs of the excavator. Head tracking of the trainee in the immersive VR is controlled by the HMD’s embedded sensors. Figure 3(c) shows the stereoscopic view that the trainee sees through the HMD.

The behaviors of agents representing the surrounding equipment are described in the previous work of the authors (Vahdatikhaki et al. 2015b). In a nutshell, the excavator moves to the first digging station and when the truck is ready, it performs the digging, swings to the truck, dumps the material, and swings back to the digging point. This cycle continuous until the truck is filled. On the other hand, the truck waits until excavators are settled in their digging stations and then moves to the excavator with the shortest queue of trucks. Then, the truck waits in the queue until the excavator is free, positions itself near the excavator, and waits until it is filled with soil. Next, it moves to the dumping point where it dumps the material, and then returns to the excavator with the shortest queue.

Once the training starts, the trainee can control the excavator to relocate to the designated digging station and declare its availability for the excavation. Once the excavator is declared ready, MAS-controlled trucks form a queue next to the excavator for the load. The trainee should then perform necessary swinging, digging and boom/stick motion to perform smooth digging and dumping operations. Once a truck is fully loaded, the truck goes to the dumping point and another truck maneuvers to be loaded by the excavator.

5. CONCLUSIONS, LIMITATIONS AND FUTURE WORK

In this paper, a novel method was presented for developing a new type of construction heavy equipment training
simulators based on the integration of the actual site data, which are collected from a wide range of sensors, and a multi-agent system in an immersive VR environment. While the presented prototype and case study are good indications of the feasibility of combining actual spatial models with a MAS system in a training VR environment, further research is required to provide a comprehensive proof of concept for the presented framework. Nevertheless, the proposed simulator is expected to offer the following advantages: (1) Representing scenarios based on the model of actual complex construction sites through which the workers can familiarize themselves with the site; (2) Providing scenarios with several pieces of equipment operated by several trainees simultaneously; (3) Capturing the realistic behavior of the surrounding equipment, where trainees can be exposed to safety-specific education; and (4) Providing multi-player immersive VR interface.

There are a number of limitations in the presented prototype and case study. First, due to the lack of relevant actual site data about equipment movement, the VR environment was developed using only the MAS. Second, the necessary parameters for feedback on safety and productivity performances, e.g., DEWs and LAEWs, are not incorporated in the presented prototype. Finally, the case study was only tested for a single trainee.

Based on the above limitations, the future work of this research will pursue the following: (1) collecting data of equipment tracking from actual sites that can be simulated in parallel to the MAS structure; (2) Improving the prototype to provide necessary feedback on safety and productivity performances as discussed in Section 3.4; and (3) Conducting a comprehensive case study with multiple users.

ACKNOWLEDGMENTS

The authors would like to thank Mr. Bruno Paes and Mr. Henrique Machado Gasparotto for their contributions to the implementation of the prototype and case study.

REFERENCES


