Joint-Space or World-Space Interfaces: Examining the Performance of Novice Excavator Operators

Ishida H., Alencastre-Miranda M., Hoshino T., Martinez Guerra, A., Usui D., Tsukahara S., and Krebs H.I., *Fellow, IEEE*

***Abstract*— We developed a hydraulic excavator simulator with a Joint-Space and a World-Space Interface in order to evaluate and train operator's skills. This simulator consists of different environments based on different construction sites and levels of difficulty in order to train and evaluate the naïve operator. The virtual excavator imitates an actual excavator's motion dynamics taking into consideration both the hydraulic system, as well as the mass and inertia of each of the structural components. To simulate the environment dynamics in real time, we employed the Unity3D physics engine to interact with solids and assets with modified shaders to simulate digging or dumping dirt. The novel World-Space Interface is a scaled version of the excavator's linkage, affording a more intuitive control to a naïve operator. Here, we report the details of the simulator, of the novel interface, and the experiments to validate this interface.**

I. Introduction

The number of skilled operators of construction machinery is decreasing in the developed countries due to aging [1]. In low and middle-income countries, there are skilled operator shortages due to a lack of training. Main reasons on why it is difficult to train operators are: 1) When training novice operators, they have to ride on the actual excavator in a training circuit; 2) The training area is artificial; 3) The excavators have a complex non-linear behavior and they are controlled via a pair of joysticks commanding the joint velocity of the 4 degrees-of-freedom (DOF). The operator must develop an internal model or map transforming the world-space requirement into the 4-DOF excavator joint-space commands. Like an infant, an operator requires years of training to achieve perfection (personal communication).

There have been some attempts to ameliorate the shortage of skilled operators. For example, Hiraoka proposed a controller that employs a command input extracted from an on-site databank storing past human commands [2]. Yuasa proposed a method to generate the excavator's bucket trajectory in terms of the amount of work and type of soil [3]. Hitachi proposed a tele-operation master-slave type controller to potentially allow one skilled operator to control two excavators [4]. Instead of the traditional bi-manual joysticks, Sun proposed a unimanual interface that would operate with only the right arm [5]. Carvalho proposed a control scheme employing an existing haptic device [6]. Semi-autonomous/autonomous operated excavators were also proposed. For example, Kobelco proposed a semi-autonomous method for simple digging motion. These efforts were limited to a single task or to multiplexing a single skilled operator [7]. Sumitomo employed machine control (MC) to trace the tip of the bucket trajectory based on drawings for slope shaping operation [8]. More recently many companies and researchers proposed similar MC concepts for semi-autonomous operation system during particular tasks such as digging, dumping, and shaping dirt. These approaches require significant effort prior to starting the actual operation. One might speculate that training of novice operators would be needed until fully autonomous excavators will be available.

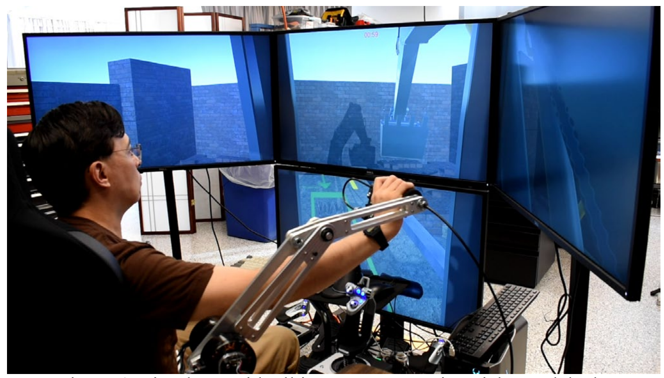

Figure 1. Simulator with all its components: joysticks, pedals, levers, four 4k monitors, high end Alienware computer, seat, and 3rd generation world space interface (which includes actuators)

In this paper, we describe our simulator for our hydraulic excavator, as well as the novel World-Space Interface (WSI) and changes in performance of novice operators while using the traditional joint-space joystick interfaces (non-linear map) and the proposed novel world-space interface (linear map) in our simulated environment.

II. Experimental Protocol

A. *Intelligent Simulator*

We developed a realistic 3D visual simulator to train and evaluate operators performing several common tasks in 19 different environments. Figure 1 shows the general diagram of the developed system that runs on a high-end computer, processing all the 3D graphics in a powerful graphics card and displaying in high resolution monitors configured as if the operator were inside the cab of an excavator. The operator can control the virtual 3D excavator using 2 Logitech X52 joysticks or a novel WSI that we designed and built. The operators can drive the virtual excavator by using the Logitech pedals or two Logitech X52 levers.

Research supported by Sumitomo Heavy Industries, Grant # 6941094.

H. Ishida, T. Hoshino, D. Usui, and S. Tsukahara are with Sumitomo Heavy Industries, Ltd., Yokosuka, Kanagawa, Japan 237-8555 (phone: +81-50-2031-5679; e-mail: hiroaki.ishida.h@shi-g.com , hoshino.takashi@shi-g.com , dotaro.usui@shi-g.com and shinichiro.tsukahara@shi-g.com).

H. I. Krebs are with the Massachusetts Institute of Technology, The 77 Lab, Cambridge, MA 02139 USA (e-mail: hikrebs@mit.edu).

A. Martinez-Guerra is presently at Noah Medical, San Carlos, CA.

M. Alencastre-Miranda is presently at Oceaneering, Austin, TX.

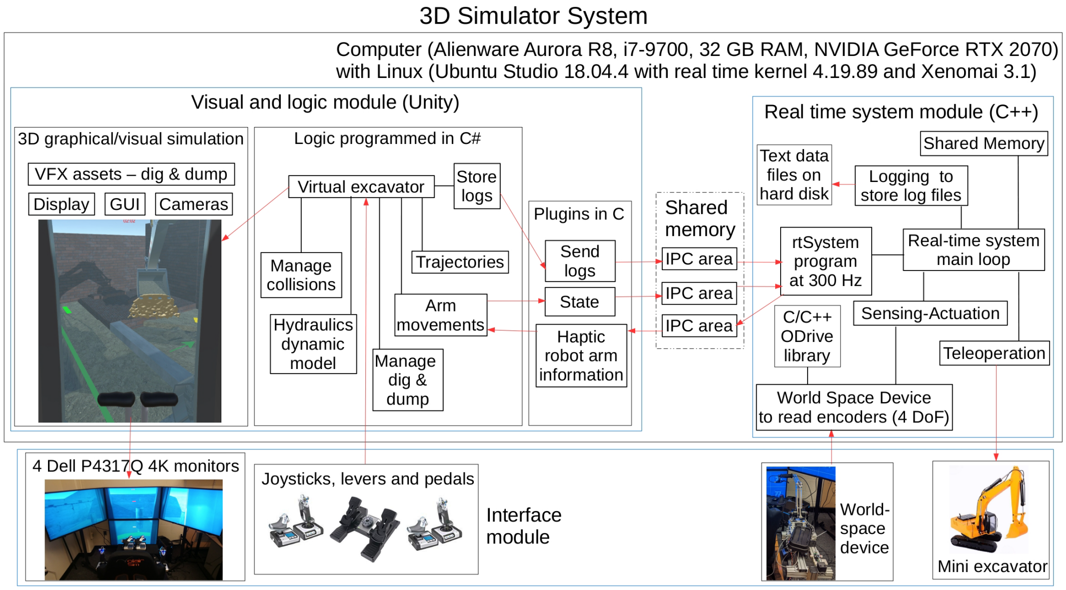

Figure 2. Diagram of the 3D simulator system

A.1 Simulator system

The whole system for the 3D excavator simulator was developed in modules. There are 3 main modules (blue rectangles) divided in blocks (black rectangles) depending on their functionality (Figure 2). The interface module includes the monitors, WSI, a mini excavator, joysticks, levers and pedals aforementioned. The real-time system module runs at 300 Hz on top of the hard real-time Linux kernel with Xenomai. One of the goals of this module is to read the encoders of the 4-DOF of the WSI and calculate the velocities sent to the simulator to control the virtual excavator. The 4-DOF correspond to the excavator cab, boom, stick, and bucket. The second goal of this module is to store the log files (time, joint angles and velocities, world space position of the bucket, etc.). The third goal is to manage the shared memory to communicate to the last two modules. The visual and logic block is the module that processes all the graphics and physics calculations to interact with the 3D simulator and display in the four 4K monitors. This last module runs at 30 frames per seconds using Unity3D. It was divided into 3 different blocks: 1) The block to render all the 3D simulation including the management of the 4 virtual cameras (for each monitor), to show the graphical user interface (GUI), and to display the visual effects (VFX) of specific assets (e.g., for the dump tasks); 2) The block where the logic of the movements and interaction of the virtual excavator are programmed counting the dynamics simulating the hydraulic system of an excavator's simple and realistic behavior, the collisions, the management for the process of dig and dump dirt, rocks or other elements in the environments, and data storage; and 3) The last block contains the plugins that manage the processes that read and write the data to the shared memory areas of the computer as part of the Inter-Process Communication (IPC) mechanism of all the real-time processes.

A.2 Training and evaluation environments.

There are 15 training environments that embody tasks such as dig and dump, remove debris, move logs or drill rocks. Those tasks are performed in different environments including construction sites, highways, roads, or a riverbed, whether in a city, mountains, forest, rocky hills or in a mine. We have also developed 4 distinct evaluation environments (not used during training) that requires an operator to fill pipes of different diameters with sand or dig a straight trench in front and dump it in the diagonal trench to the right and flatten it or to make a slope up (65 deg) and down (10 deg).

Of notice, we used Unity3D Physics Engine as a proxy to deal with the dynamics of moving solid debris, trunks, pipes, etc in some environments. In other environments, we used a Unity asset, originally used to destroy elements in games, to simulate the cracking and breaking of stones or granite slabs (e.g., using a jack hammer). For the digging, we modified a Unity asset to simulate snow through deformable grounds using compute shaders. For the dumping task, we modified another Unity asset to simulate fluids using shaders and particles to generate the visual effect of dumping the corresponding material from the bucket. We also use a more sophisticated Unity asset (based on voxels and Marching Cubes) to modify the terrain as the operator is digging or dumping dirt. This means that we are modifying the mesh to decrease its height when a volume of soil, from a 3D area defined of the terrain, is removed, and to increase its height above another area where we are dumping the dirt. We have not yet compared the quantitative performance of expert operators in our simulator with their performance in a real excavator, but we ran a questionnaire among 9 Sumitomo expert operators and at least on a qualitative level our simulator is very realistic. We ran the questionnaire during the fine-tuning phase (when we adjusted the kinematic and dynamic parameters of the simulator).

A.3 World-Space Interface

We designed a novel "World Space Interface" (WSI) shown in Figure 3. WSI requires naïve operators to linearly scale the movements required at the end-effector. In other words, WSI eliminates the need of building an internal non-linear map from joint-space to world-space (US Provisional Application No.: 63/481,021, 23 Jan 2023 and PCT/US2024/012469, 22 Jan 2024). The WSI is a scaled version and follows the revolutional joint configuration of a conventional excavator (Figures 4). When selecting the dimensions, we considered male anthropometric ergonomics, so that WSI has sufficient range to allow most men to complete the operation, resulting in L_1 and L_2 link lengh of 240[mm] and L_3.link of 46[mm]. We placed 2 springs, K_1 and K_2, in the parallel linkage mechanism to provide anti-gravity support including the weight of the parallel linkage, the brake, and the average man's arm weight (~3.5[kg]) [9]. The goal is to reduce the operator's fatigue. The WSI has rotary and linear encoders. These signals are sent to the computer via serial communication at 300[Hz].

B. Training and Evaluation Protocol

B.1 Naïve Subjects

12 naïve subjects participated in our experiment. These subjects were healthy adult males, 24 to 35 years old (average 29 y.o.), who have never had any experience operating actual excavators. We divided them into 2 groups; one trained only with WSI and another only with the joystick interface (JS).

B.2 Protocol Description

For training, 14 different environments requiring subjects to learn different skills, such as digging, dumping, drilling, etc. Each task was designed so that the operators could finish it in ~30 minutes. Each subject was asked to operate the excavator in 2 different environments each day (total of 7 days).

We employed the 4 evaluation environments to assess operator's skills. The operators dig the sand in the circle around the excavators and dump it into 8 pipes (Figure 5). The diameter of each pipe gradually decreases and the difficulty of positioning the bucket tip increases. In evaluation environment 2, simultaneous movements of the cab and other joints were evaluated. Experimental subjects made 4 trenches in front of the excavator and dumped the sand into 4 different trenches, which were arranged diagonally, and then flattened the surface of these diagonal trenches. This operation with the joysticks along the diagonal lines is quite difficult. In evaluation environments 3 & 4, fine motor coordination was measured. The subjects remove the sand along the specified slopes: 4 up and 3 down slopes with different inclination angles. These movements are also considered to be difficult with the joystick interface.

The day prior to the initial training day and following completion of the training, subjects used the unused training environment for 30 minutes for acclimatization (we built 15 enviroments and used 14 during training). Then subjects were requested to finish the 4 evaluation tasks as fast as possible. The comparison of the performance before and after the training can assess the training's effectiveness.

B.3 Description of Metrics

In this paper, we report on 6 different metrics to evaluate the operation skills: 1) cycle time, 2) tip distance error when dumping, 3) number of joints operated simultaneously, 4) idle time ratio 5) task speed, and 6) performance in Fitts's law in the evaluation environment No.1. Cycle time is defined as time duration from the beginning of the digging movement to the completion of the dumping movement for each repetitive attempt. Tip distance error at dumping is the horizontal distance between the bucket tip and the center of each pipe when the dumping motion starts. The idle time ratio is a percentage of the cycle time that the speed is less than 10% of maximum achievable joint speed. The task speed indicates how fast the sand height in the pipe changes.

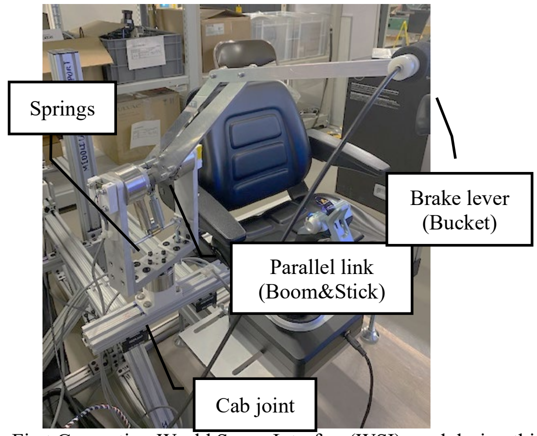

Figure 3. First Generation World Space Interface(WSI) used during this initial test with subjects which does not include actuators. Newer versions of the WSI include actuators to afford haptic experience.

Figure 4. Mechanism and coordination of the simulated excavator

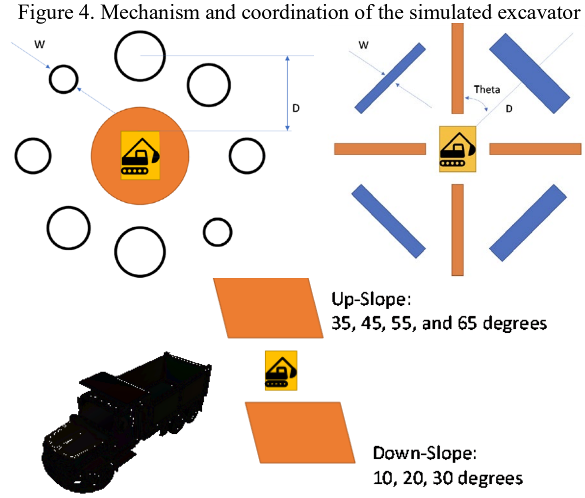

Figure 5. Evaluation No.1 (top row left), 2 (top row right), 3&4 (bottom).

The number of joints indicates the average number of excavator joints which are operated simultaneously. The performance metric b is defined as $\frac{1}{b} = \frac{\log_2 \frac{2D}{W}}{t}$ where D is the distance from the center of the digging area to the center of the pipe and equal to 3.263[m] and W is diameter of the pipes ranging from 0.8, 0.9. 1.0, 1.2[m], t is time to carry sand from the digging area to the pipes. This performance index b represents the learning rate in Fitts' law; $T = a + b\log_2(1 + D/W)$. Failures, i.e., hitting obstacles, were eliminated prior to any analysis.

III. Results

In Figure 6, we show the average time of the aggregate of the group using WSI and JS pre-and-post training. The subjects with JS can improve and decrease their cycle time with training by 46%, the cycle time with WSI does not change much (1% increase). A simple t-test from pre-to-post training leads to a non-statistically significant p-value of 0.94 for the WSI. The p-value when using the JS is 0.019. The other 5 metrics show similar results except for the task speed (i.e., how fast sand is dropped into the pipes [1/sec], p=0.027). Comparing the pre-training results with the JS and WSI, the tip distance errors is significantly different (p=0.041), and the accuracy with WSI is 17% higher than with JS after training. For the number of joints, we observed significant difference between WSI and JS at pre-training (p=0.009), task speed (p=0.005), and for the performance index (p=0.025) with a significant pre-to-post training difference with JS for the idle time ratio (=0.042), task speed (p=0.006), and performance index (p=0.004).

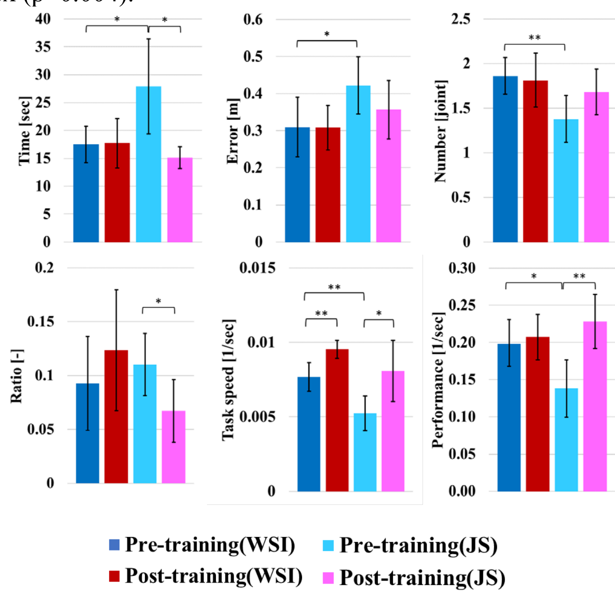

Figure 6. Comparison of each metric results for 12 naive subjects with 6 of them using the WSI and the other 6 novices using the joysticks (JS) pre-to-post training (*p < 0.05 for significance).

IV. Discussion

First, we demonstrated that operating the excavator with the WSI is quite intuitive. From the comparison of the cycle time before the training, the average time with WSI is 37% shorter than with JS. Furthermore, the time with WSI before the training is almost the same as the cycle time post-training with either interface. Same trends can be seen in other metrics.

The second point is the training effectiveness with WSI. The average results pre-to-post training indicate that the naïve operator does not improve his operational skills much with WSI, perhaps because the task is not challenging enough.

Third, the pre-training average result in 6 metrics with WSI is close to the results with JS post-training. This result indicates that WSI can potentially provide high operability at first use and could be implemented in actual excavators to allow naïve operators to start working immediately.

It is worth mentioning that the results are inline with human motor control and it is possible that artificially increasing the *visual* consequences of the deviations of reaching movements with the WSI in the simulator might improve adaptation and learning.

Lastly, our simulator still requires improvement. The shaders did work well to simulate sand or dirt but we want to improve further as we will explore whether haptic feedback will facilitate learning. To that effect, we are developing an "in-house" smooth particle hydrodynamics (SPH) method to simulate the soil interaction. We prefer SPH over the commonly adopted DEM methods (discrete element model) for real time applications with our haptic joysticks and WSI.

V. Conclusion

We proposed a hydraulic excavator's 3D simulator with JS or a novel WSI. We validated the effectiveness of WSI when operated by naïve subjects with the performance metrics indicating that the pre-training performance with the WSI is similar to the performance with JS post-training and it does not improve with further training. Inline with our hypothesis that from the onset the WSI can provide the naïve operator with an intuitive way to operate excavators.

While expert operators believe the simulator can be used for training and, at least qualitatively, closely resemble the "real world," we will augment it and investigate further ways to speed up training. In particular, we want to understand the impact of haptic feedback on both interfaces, i.e., JS and WSI.

Acknowledgment

We thank all subjects from the Technology Research Center at Sumitomo Heavy Industries, Ltd. and Sumitomo Construction Machinery for participating in our experiments. MIT-COUHES (Committee on the Use of Humans as Experimental Subjects) deemed the protocol exempt (E-3460) under Exempt Category #3 - Benign Behavioral Intervention.

References

[1] Marsh & GuyCarpenter, "Future of Construction -A Global Forecast Construction to 2030," Oxford Economics, 2021.

[2] K. Hiraoka, T. Yamamoto, M. Kozui, K. Koiwai, K. Yamashita, "Design of a Database-Driven Assist Control for a Hydraulic Excavator Considering Human Operation," J Rob Mech, 35:3, 703-710, 2023.

[3] T. Yuasa, M. Ishikawa, S. Ogawa, "An Optimal Design Methodology for the Trajectory of Hydraulic Excavators Based on Genetic Algorithm," Journal of Robotics andMechatronics, 33:6, 2021.

[4] Hitachi,https://www.hitachicm.com/global/businesses/products/double-arm-working-machine/, Feb. 22th. 2022.

[5] D. Sun, S. Lee, Y. Lee, S. Kim, J. Ueda, Y. Cho, Y. Ahn, C. Han, "Assessments of Intuition and Efficiency: Remote Control of the End Point of Excavator in Operational Space by Using One Wrist," ASCE Int Conf on Computing in Civil Engineering, 273-280, 2019.

[6] L. Carvalho, P. Rezeck, M. Lima, L. Pinto, G. Freitas, E. Nascimento, D. Macharet, L. Chaimowicz, G. Pessin, M. Campos, "On the Evaluation of Force Feedback Augmented Teleoperation of Excavator-like Mobile Manipulators," 2020 IEEE 16th Int Conf on Automation Science and Engineering (CASE), 1401-1407, 2020.

[7] M. Takeishi, T. Doi, D. Noda, "On-site demonstration of automatic operation hydraulic excavator," 2020 Symposium on Construction Machinery and Methods, 189-194, 2020. (in Japanese)

[8] Sumitomo, https://www.sumitomokenki.com/, Feb. 22th. 2022.

[9] N. Diffrient, A. R. Tilley, J. C. Bardagjy, "Humanscale 1/2/3. 4/5/6, 7/8/9," 1974.