2012 Third International Conference on Intelligent Control and Information Processing
July 15-17, 2012 - Dalian, China

Learning and Control in Virtual Reality for Machine Intelligence

Xiao Fang, Haibo He, Zhen Ni and Yufei Tang

***Abstract*— This paper presents a Virtual Reality (VR) interactive platform for learning and control for machine intelligence based on Adaptive Dynamic Programming (ADP). Recent research results have provided strong evidences that ADP could be a key technique for brain-like intelligent systems design, and VR is a powerful human-computer interface which can provide a three-dimensional (3D) active virtual environment. Converge these two subjects, we design an interactive system to facilitate and demonstrate the learning and control in VR environment with ADP. Our main goal in this paper is two-fold. First, we demonstrate that VR could be a useful platform to demonstrate and visualize machine intelligence research through the simulated 3D environment. Second, the integration of VR into machine intelligence research can provide a powerful platform to simulate, validate and facilitate the real-time interaction between the intelligent system and an unstructured environment. We discuss the detailed design strategy of the VR platform, and also demonstrated the interactive system performance based on the triple link inverted pendulum benchmark.**

I. Introduction

THE stabilization of nonlinear systems has been extensively studied for learning and control research purposes [1]. Besides experimental testing with the real hardware control system, an effective approach is to design a platform to simulate the dynamic progressing and system behavior under a simulated virtual environment [2]. The simulation process demonstrate experiments with the virtual model of a real system for the purpose either of understanding the behavior of the system or of evaluating various strategies for the operation of the system [3]. This is particularly important to understand the interactions among an intelligent agent, learning and control approach, and the external environment. There are two key techniques in this aspect: computation and visualization [4]. In this paper, we aim to develop an integrated VR environment with the ADP controller and demonstrate its effectiveness in support of machine intelligence research.

VR is a computer-controlled multisensory communication technology, which provides a new way of intuitive interaction between data and human decision [5]. The types of VR devices could be classified as head-mounted display (HMD), fiber-optic wired gloves, position tracking devices, desktop virtual reality, and other [6]. Recently, the most exciting aspect in VR research is not the advancements in each specific device, but the increasing adoption of its technologies combined with other scientific fields [5]. For this reason, VR technology has attracted growing attention in different disciplines.

X. Fang, H. He, Z. Ni and Y. Tang are with the Department of Electrical, Computer and Biomedical Engineering, University of Rhode Island, Kingston, RI 02881, USA (email: {xfang, he, ni, ytang}@ele.uri.edu).

This work was supported in part by the National Science Foundation under Grant ECCS 1053717 and the K. C. Wong Education Foundation, Hong Kong.

In this paper, we are interested in VR for machine intelligence research through real-time interaction of the control system with the external environment. The main contribution is that we develop an integrated platform that can demonstrate and visualize complex system behaviors under different scenarios without the requirement of setting up expensive or lab-intensive physical experiments [7]. In addition to its unique visualization capability, our approach also provides the human-in-the-loop interaction to facilitate human-machine interface.

In this work, we combine the VR platform with learning and control system. We use the ADP algorithm to solve the optimal control problem and use the VR platform to imitate the real environment, to simulate the real-time controlling progress, to verify our learning method intuitionally and to interact with our controller system.

The rest of the paper is organized as follows. In section II, we present the overall architecture of the interactive VR system and introduce each module in detail. In section III, we test the interactive system on triple link inverted pendulum balancing problem. Finally, a conclusion is given in section IV.

II. Overall System-level Architecture of VR with ADP

To develop the interactive system, we first design a closed-loop control circuit that the system with ADP. A VR environment is integrated into this system as a real-time simulation environment, which could animate the simulation progress and exchange data between the environment and system-under-control. To do this, we design a 3D virtual system which could input the state variables from the system and read these variables to change the status of simulated object such as position, rotation angle, speed, among others. Finally, we design a feedback module which we could change parameters of the virtual object in real-time simulation process. Fig. 1 shows the overall structural of our interactive VR system.

A. Design of the ADP Module

ADP has become a key method for adaptive control and learning for nonlinear systems [8][9][10]. To approximate the cost function J, ADP is built on Bellman equation using a function approximation structure such as neural networks to achieve optimization over time [9]. This can be accomplished by adapting the weights of the critic network to make the approximating function J [11]. Generally, given a system with a performance cost function, the objective of dynamic

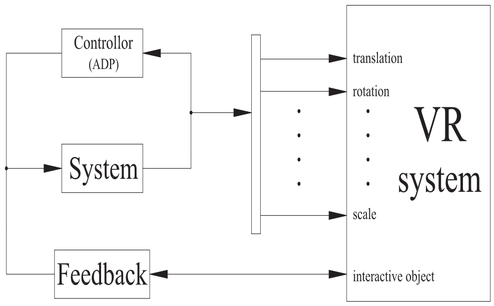

Fig. 1. The over architecture of the interactive VR system with ADP

programming is to choose a control sequence $u(k)$ so that the cost function is minimized

$$J^*(x(k)) = \min_{u(k)}\{U(x(k), u(k)) + \alpha J^*(x(k+1))\} \quad (1)$$

where $x(k)$ is the state vector of the system, $u(k)$ is the control action, U is the utility function, and α is a discount factor. In this way, ADP can successfully achieve learning and control by using a functional approaching structure to approximate the cost function. Among various categories of ADP method, heuristic dynamic programming (HDP) is the most straightforward implementation of ADP design. In HDP, the critic network estimates the total cost function with local cost/utility function value U [9]. In this paper, we adopt the action dependent heuristic dynamic programming (ADHDP) approach as proposed in [11] as our ADP control approach, in which we store the previous J value instead of using a model network to predict the future system states and consequently the cost-to-go for the next time step [11]. For details about this ADP structure, interested readers can refer to [11] for further details.

B. Design of the VR Module

VR module is a virtual platform for simulating the dynamics of systems. To design it, first we need to develop a 3D environment to mimic the real system. There are numerous ways for designing such a virtual environment with a realistic performance [12]. In this paper, we design the 3D object model by using Auto CAD and Auto 3ds Max software, and then export the system in VRML (Virtual Reality Modeling Language) format. The system can be interpreted by VRML explorer in three-dimensional scene in order to be integrated with the Matlab Virtual Reality environment for modeling and simulation [13].

Furthermore, to enable human-computer interaction through the VR environment, we developed a “VR text block” in our virtual environment, which takes input vector using the format string defined in its mask and the resulting string is sent to the “string” field of the associated VRML text node in the scene [14]. In this way, the input state variables can dynamically change the VR object status to provide active human interaction during simulation.

C. Design of the Feedback Module

During simulation, the parameter of 3D object properties in the virtual scene can also be read into Simulink [12].In VR technology, VRML is a standard language for describing 3D image sequences and possible user interactions to go with them. It provides many interactive sensor nodes to simulate the verisimilar virtual environment such as touch sensor, visibility sensor, proximity sensor, collision sensor and so on. In this paper, we use the touch sensor to construct the interactive system. The touch sensor reacts when the user has the mouse over or when the user clicks a shape area in VR environment. According to this value, we create an object on touch area which could be changed its position by user click mouse on touch area. Then we output the trajectory of this object into feedback module. The key of the feedback module is a Trajectory Graph block, this block is embedded by S-function which could compute the value of the trajectory and output this value to system as an external disturbance. This disturbance could break the control system’s equilibrium and lead the ADP controller re-learn to the balance point.

III. System Design in VR Environment

As the triple link inverted pendulum balancing task is a popular benchmark in the community [15], we use that benchmark to validate and demonstrate our VR learning and control platform. To do that, we first discuss the detailed design strategy of implementing the system in 3D virtual environment.

A. Physical Model

Fig.2 shows a schematic of the triple link inverted pendulum, with three links of variable length mounted on a cart. Here the length of each link $L1$, $L2$, $L3$ is 0.43m, 0.33m, and 0.13m, respectively. The cart is free to move within the bounds of a one-dimensional track while the links are free to rotate only in the vertical plane of the cart and track [15][16].

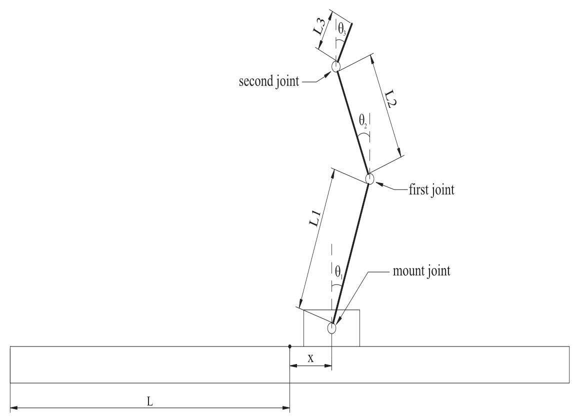

Fig. 2. Schematic of the triple link inverted pendulum

The equation governing the system can be described as follows:

$$F(q)\frac{d^2q}{dt^2} = -G(q, \frac{dq}{dt})\frac{dq}{dt} - H(q) + L(q,u) \qquad (2)$$

From the above nonlinear dynamical equation, the state-space model can be described as follows:

$$\dot{Q}(t) = f(Q(t), u(t)) \qquad (3)$$

with

$$\begin{aligned} & f(Q((t), u(t)) \\ & = \begin{bmatrix} 0_{4\times4} & I_{4\times4} \\ 0_{4\times4} & -F^{-1}(Q(t))G(Q(t)) \end{bmatrix} Q(t) \\ & + \begin{bmatrix} 0_{4\times4} \\ -F^{-1}(Q(t))[H(Q(t)) - L(Q(t), u(t))] \end{bmatrix} \end{aligned} \qquad (4)$$

and

$$\begin{aligned} Q(t) = [\ & x(t) \quad \theta_1(t) \quad \theta_2(t) \quad \theta_3(t) \\ & \dot{x}(t) \quad \dot{\theta}_1(t) \quad \dot{\theta}_2(t) \quad \dot{\theta}_3(t)\ ]^T \end{aligned} \qquad (5)$$

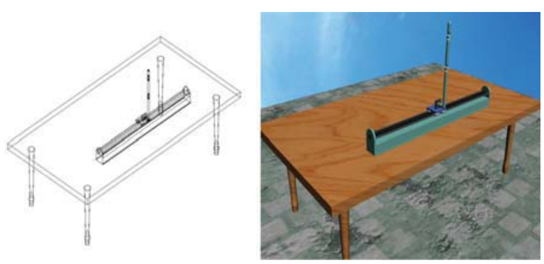

Fig. 3. Development of the 3D environment

The output of eight state variables are: 1) $x(t)$, position of the cart on the track; 2) $\theta_1(t)$, vertical angle of the first link joint to the cart; 3) $\theta_2(t)$, vertical angle of the second link joint to the first link; 4) $\theta_3(t)$, vertical angle of the third link joint to the second link; 5) $\dot{x}(t)$, cart velocity; 6) $\dot{\theta}_1(t)$, angular velocity of $\theta_1(t)$; 7) $\dot{\theta}_2(t)$, angular velocity of $\theta_2(t)$; and 8) $\dot{\theta}_3(t)$, angular velocity of $\theta_3(t)$[11].

In this paper, the constraints for this system include the following: 1) the cart track extends L to both ends from the center position is 1.0 m; 2) each link angle should be within the range of $[-20°, 20°]$ with respect to the vertical axis [1][11].

B. Design of the 3D Model

The 3D model is developed by using Auto CAD and 3ds Max according to the requirements of the schematic model, and the constraint conditions are coded using VRML. Fig.3 demonstrates a snapshot of our design. After the 3D model is converted into a vrml file, it is then imported into Matlab Simulink 3D Animation system for simulation in VR environment.

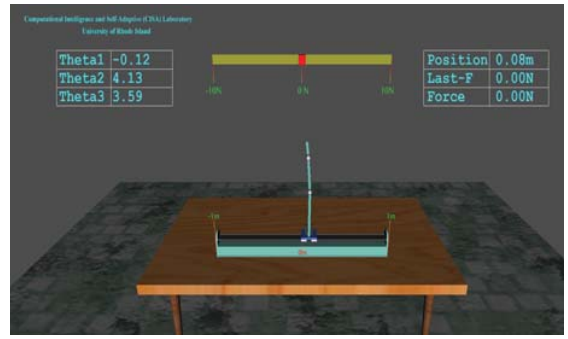

Fig. 5. A snapshot of the developed VR environment for learning and control

C. System Integration with in VR Environment

To simulate the system dynamics, we design a closed loop structure to connect the ADP controller and benchmark system. The eight variables are communicated in this closed loop system. At the same time, four of these variables $(x(t), \theta_1(t), \theta_2(t), \theta_3(t))$ will be taken as inputs to VR module to visualize the real-time animation of the dynamic control process. Besides simulation, we also develop a "F-box" model and a touch sensor area in VR environment that could be used as an external input to facilitate human-computer interface. This means during the real-time simulation process, we could click the mouse on touch sensor area to change the X axis coordinate of "F-box". We record the X axis coordinate of "F-box" to provide an external input (in our current case, the applied force) to the system. The trajectory of "F-box" will be computed by Trajectory Graphs module and output to system as an impulse signal. This impulse signal could simulate the realistic "left" or "right" force on cart in VR environment. Fig. 4 demonstrates the schematic diagram of triple link inverted pendulum interactive system in Matlab/Simulink. Based on this, we present the simulations on Matlab/Simulink to verify this interactive system, as demonstrated in Fig.5.

D. Simulation Results and Analysis

We present the simulation results to verify and demonstrate the learning and control capability of our system in VR environment. Fig. 6 shows the simulation results for the four variables, the position of cart and the vertical angle of three links. From this figure one can see, the inverted pendulum reaches the balance condition after about 350 time steps.

In order to test how the ADP controller is working under the condition of external disturbance, we add a virtual force to the cart in the VR world after 780 time steps by changing the position of the "F-box". This force is added to the system as a pulse signal to break the system's equilibrium. From the system states in Fig.7, one can clearly see at 780 time step the system will quickly response to this external force. The ADP controller will observe the new system states and adjust its control action accordingly to balance the triple link inverted pendulum. This is clearly shown in this figure by

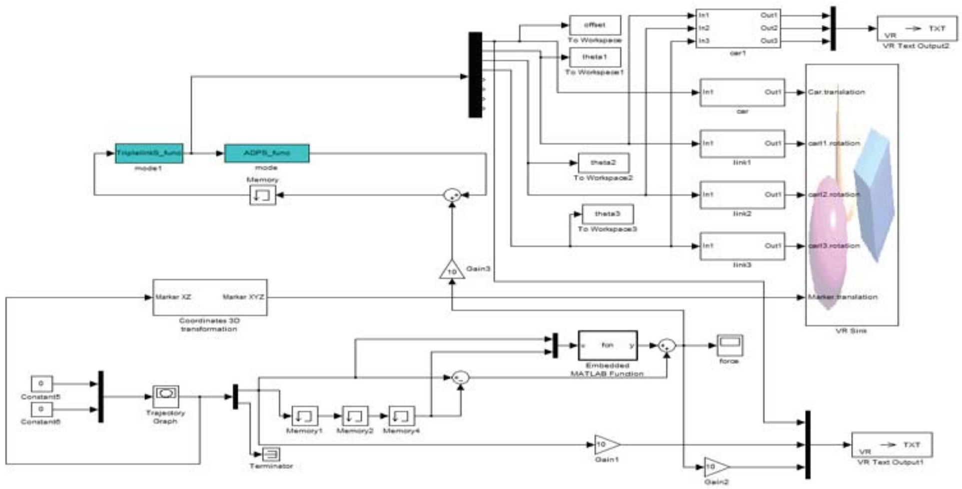

Fig. 4. Schematic of the triple link inverted pendulum system with ADP controller and VR system in Matlab/Simulink

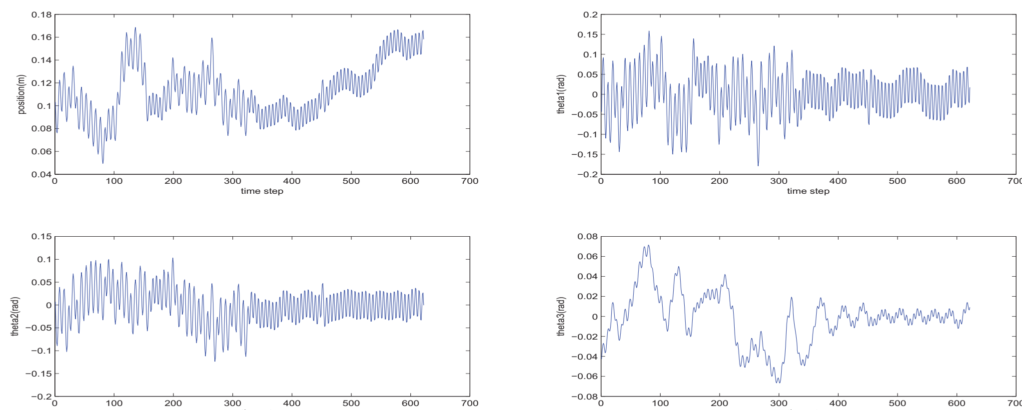

Fig. 6. Simulation results of the triple link inverted pendulum system

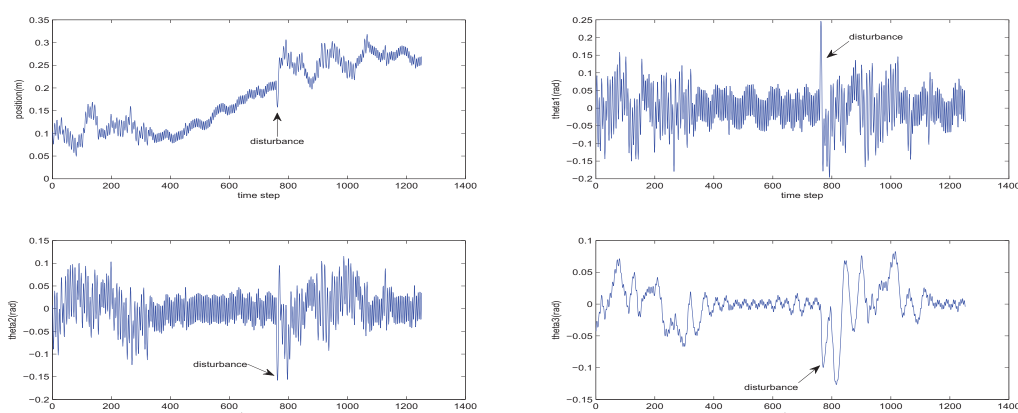

Fig. 7. Simulation results with adding external virtual force at 780 time step

observing the state vectors. This suggests the ADP approach in this case is robust and can adaptively control the system with relatively large external disturbance. This also suggests that our VR platform is an effective approach to design and test different scenarios to facilitate the machine intelligence research.

IV. Conclusion

In this paper, we develop a learning and control system in VR environment with the integrated ADP controller. We present the system architecture with detailed design strategies of each module. To verify the effectiveness of our approach, we use the triple link inverted pendulum as a benchmark to demonstrate the learning and control capability in a virtual environment. Furthermore, our VR platform enables external human interaction with the system-under-control, which is an important capability to verify and facilitate machine intelligence research with human-in-the-loop. We hope this VR environment could provide a useful platform for advanced control and learning approach developments to facilitate the machine intelligence research.

References

[1] K. D. Eltohamy and C.-Y. Kuo, “Nonlinear optimal control of a triple link inverted pendulum with single control input”, *Int. J. Contr.*, vol. 69, No. 2, pp. 239-256, 1998.

[2] J. T. T. Chuan and F. H. Nagi, “Modeling and simulation of differential drive mobile robotics system”, *Mobile Robotics Group, University Tenaga Nasional*

[3] Ricki G. Ingalls, “Introduction to simulation”, *Proceedings of 2002 Winter Simulation Conference*, pp. 7-16, 2002

[4] Nobuaki Ohno and Akira Kageyama, “Introduction to virtual reality visualization by the CAVE system”, *Advanced Methods for Space Simulations*, pp. 167-207, 2007

[5] Hilary McLellan, “Virtual reality: a selected bibliography”,*Educational Technology Publications*, 1992

[6] G. Burdea and P. Coiffet,“Virtual reality technology”, *2nd Edition, Wiley-Interscience*, 2003.

[7] Simulink 3D animation, *http://www.mathworks.com/products/datasheets/pdf/simulink-3d-animation.pdf/*

[8] Jian Fu, Haibo He and Xinmin Zhou, “Adaptive learning and control for MIMO system based on adaptive dynamic programming”, *IEEE Transacitons on Neural Networks*, vol. 22, No. 7, pp.1133-1148, July 2011

[9] Feiyue Wang,Huaguang Zhang and Derong Liu, “Adaptive dynamic programming: an introduction”, *Computational Intelligence Magazine*, pp. 39-47, 2009

[10] Haibo He, Zhen Ni, and Jian Fu, “A three-network architecture for on-line learning and optimization based on adaptive dynamic programming”,*Neurocomputing*,vol.78, pp.3-13, 2012.

[11] Jennie Si and Yu-Tsung Wang, “On-line learning control by association and reinforcement”,*Transactions on Neural Networks*,vol.12, No.2, pp. 264-276, March, 2001

[12] Daniel Dominguez, “VRML and simulink interface for the development of 3D simulator for mobile robots”, *World Academy of Science, Engineering and Technology*, pp. 281-283, 2007

[13] Jozef Novak-Marcincin, “Application of the virtual reality modeling language for design of automated workplaces” , *World Academy of Science, Engineering and Technology*, 31, pp. 160-163, 2007

[14] Simulink 3D animation - plane take-off with HUD text, *http://www.mathworks.com/products/3d-animation/*

[15] K.G. Eltohamy, C.Y. Kuo.“Real time stabilisation of a triple link inverted pendulum using single control input.” *IEEE Proc.-Control Theory Appl*, vol. 144, No.5, pp.498-504, 1997

[16] Andrew G. Barto, Richard S. Sutton and Charles W. Anderson.“Neuronlike adaptive elements that can solve difficult learning control probelms.” *IEEE Transactions on Systems, Man and Cybernetics*, vol. SMC-13,No.5, pp.834-846,1983