IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES

ISSN PRINT 2319 1775 Online 2320 7876

Research paper  Volume 11, Iss 4, 2022

# REMOTE EXPERMENTATIONS OF ARTIFICIAL INTELLIGENCE IN EDUCATION

**[1]Dr. R. Rambabu, [2]Dr. D. Naga Purnima, [3]G. Swarnalatha**

*[1]Professor & HOD, Department of Computer Science & Engineering, Rajamahendri Institute of Engineering & Technology, Rajamahendravaram, A.P, India*
*[2]Professor& HOD, Department of Mathematics, Rajamahendri Institute of Engineering & Technology, Rajamahendravaram, A.P, India*
*[3]Associate Professor, Department of Computer Science & Engineering, Rajamahendri Institute of Engineering & Technology, Rajamahendravaram, A.P, India*

## ABSTRACT:

Research on the application of visual artificial intelligence (AI) in education has advanced significantly in recent years. should take advantage of the opportunity to gather a lot of data in many settings and circumstances. However, gathering this kind of data requires a lot of work and time. In addition, creating and evaluating visual AI algorithms for multisensory models are costly and sometimes hazardous real-world procedures. To solve both of these problems, a 3D environment simulator that offers variable setup of multimodal sensors and produces photo-realistic simulations using a view synthesis module. To produce realistic images, we incorporate innovative depth refinement, adaptive view selection, and layered rendering into our view synthesis module. It demonstrates the various benefits that PreSim offers: Three key features it offers are: (i) a photo-realistic 3D environment that makes it easy to integrate multisensory models in the virtual world and allows them to perceive and navigate scenes; (ii) an internal view synthesis module that makes it possible to translate simulation-tested algorithms to physical platforms without domain adaptation; and (iii) the capacity to generate large amounts of data for vision-based applications, like object pose and depth estimation. Thus, students can profit from virtual classrooms by adopting.

**KEYWORDS:** Simulation and Animation, Sensor Fusion, RGB-D Perception, Remote Experimentation, 3D virtual worlds

## I. INTRODUCTION

Deep network-based data-driven approaches have shown remarkable performance in recent years for computer vision problems including 6D object pose estimation and depth estimation [1]. A lot of data is required for these data-driven techniques to train and evaluate their models. But gathering and classifying data is a laborious and time-consuming task. The simulated environment is starting to show promise as a solution to these issues since it can supply a large amount of annotated data for a variety of AI activities.

A major current focus of environment simulators is to reproduce high-quality free-viewpoint rendering of real scenes. There are a number of open source simulators [2] to achieve this goal by parameter settings of scene details, including geometry, texture, lighting and 3D modeling of static objects. However, parameter setting is time-consuming and labor-intensive. Even with precise modeling and suitable parameter settings, the simulated world still lacks richness and diversity of the real world. This disadvantage may result in the failure of transferring algorithms that are developed and tested in simulation to physical platforms for many vision-based tasks, such as object recognition, obstacle avoidance, and visual navigation. This problem is known as the reality gap: the discrepancy between synthetic and real data.

To address this issue, game engines which allow photorealistic rendering have been leveraged to build virtual environments.

ISSN PRINT 2319 1775 Online 2320 7876
Research paper  Volume 11, Iss 4, 2022

However, the simulated environment heavily depends on the game engine's detailed datasets, which makes it impossible for users to build their own environments with their own datasets. On the other hand, game engines often use 3D graphics pipelines to provide real-time rendering. Thus, the rendering time increases linearly with the number of polygons to be rendered (scene complexity). To achieve real-time performance, it requires dedicated hardware and architecture design for 3D graphics. On the contrary, image based rendering which can provide real-time realistic imagery does not have these limitations [3]. It only requires a sparse collection of captured images and allows a 3D scene to be visualized realistically without full 3D reconstruction. This approach has shown high-quality results in various environments. In addition, the run time of image based rendering mainly depends on the display resolution of the output image rather than scene complexity.

Taking advantage of image based rendering, we introduce PreSim which is a 3D photo-realistic environment simulator for training and testing AI algorithm. To narrow the reality gap between simulation and reality by providing huge amounts of photo-realistic virtual RGB-D views from arbitrary locations for vision-based applications. The main contributions of our simulator are: Aphoto-realistic 3D virtual environment that provides users with ground truth poses of the multisensory model and free viewpoint color-and-depth image pairs, even in regions where a global 3D reconstruction of the scene has inaccurate or missing data. A global visualizes providing real-time positions and whole trajectories of moving sensors, and a global 3D map. A sequence controller and recorder components to control the movement of sensors and store all the required information for developing AI algorithms [4]. A novel view synthesis module built on image based rendering that combines depth refinement, adaptive view selection and layered 3D warping to lower the rendering complexity and improve the quality of synthesized images.

An entertainment applications, artificial intelligence techniques have most often been used to implement embodied agents or to automatically generate artistic content. A more recent development concerns using AI to support the user experience through new AI-based interactivity techniques. This is especially of interest for the development of artistic installations based on interactive 3D worlds [5]. A major difficulty in developing such installations is to properly translate the artistic intention into actual elements of interactivity, which in turn determine the user experience. The starting point of this research was to facilitate the description of high-level behaviors for virtual worlds that would form part of Virtual Reality (VR) art installations. Our underlying hypothesis has been that AI representations inspired by planning formalisms, and AI-based simulation derived from these, could constitute the basis for virtual-world behavior in these installations.

In our approach to interactivity, the consequences of user interaction can be dynamically computed to produce cascaded effects eliciting a specific kind of user experience. This chain of events is computed from first principles embedding elements of the artistic brief (the artist's initial conceptual description of the interactive installation and the intended user experience). In other words, AI techniques are used for their ability to represent actions and to compute analogical transformations on them to create a user experience [6].

In addition, what makes possible the use of AI techniques to simulate behavior in virtual worlds is the exploitation of a specific feature of game engines, namely

ISSN PRINT 2319 1775 Online 2320 7876
Research paper  Volume 11, Iss 4, 2022

the fact that they rely on event-based systems to represent all kinds of interaction. Event based systems originated from the need to discretize physical interaction to simplify physical calculations: although the dynamics of moving objects would be subject to numerical simulation, certain physical interactions' consequences (for example, glass shattering following impact from a hard object) could be determined in a discretized system without having to perform complex mechanical simulations in real time. Our installation consists of a virtual world in which we can alter the normal laws of physical causality by substituting physical actions' default effects with new chains of events.

## II. LITERATURE SURVEY

A. De Lucia, et al. [7] presents a virtual campus created using Second Life which provides four distinct types of virtual space: common student campus, collaborative zones, lecture rooms and recreational areas, they argued that, in a 3D multi-user virtual environment, learning is strongly related to the user perception of belonging to a learning community, as well as to the perception of awareness, presence and communication. They conducted an experiment involving university students aiming at evaluating Second Life synchronous distance lectures in the proposed learning environment which results are very positive.

D.C. Cliburn and J.L. Gross, et.al [8] used a quasi-experimental pretest-posttest comparison groups design to compare the experience of a Second Life lecture to a real world lecture, found that those who attended the real world lecture performed significantly better on a posttest quiz than those who attended the same lecture in Second Life., commented that students encountered many difficulties, such as problems viewing the lecture material, and a lack of constraints on avatar behavior in the educational setting.

C. Hong et.al [9] described the modern service industries highly depend on new technologies and skills of service talent to carry out business in a flattened world, services skills are not like basic concept or knowledge of science and technology. They could not be very easily trained in classroom without practice. Service skills are often consist of soft skills especially those skills that require a service staff to interact with relevant stakeholders, to be aware of the service environment, to respond to multiple groups' requests based on candidate solutions, priority of tasks, and all agreed final decisions.

P. Dev, et al [10] reported a project of developing and evaluating a computer-based simulator (the Virtual Emergency Department) for distance training in emergency medicine residency programs teamwork and leadership in trauma management which aimed not only to manage trauma effectively but also not needing practice on live patients.

S. Bronack, et al [11] utilized a social constructivist framework analysizing 3D virtual world learning environment is unlike traditional classroom- or web-based learning environments in many important ways, students should be provided more choices within the 3-dimensional world, should be aided to construct individual paths through the virtual world .

L. Jarmon, et al [12] suggests that 3D virtual worlds can be well suited for experiential learning environments, they use mixed research methods of journal content analysis, surveys, focus group, and virtual world snapshots and video, empirically examines the actual instructional effectiveness of Second Life as an experiential learning environment for interdisciplinary communication.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper  Volume 11, Iss 4, 2022

R. Ortiz-Cayon, A. Djelouah, and G. Drettakis et al. [13] divide the image into super pixels to preserve object boundaries and then project each super pixel to the virtual view by a local shape-preserving warping to improve the blending quality. However, this approach does not consider photo-consistency and still suffers from silhouette flattening and inaccurate occlusion edges. There have been several works that improve the quality of synthesized images by filling holes. Nevertheless, the numbers of input views are fixed in these methods, which may lead to hole filling failure when the chosen views are useless or redundant. Instead, we use an adaptive view selection method to avoid such case.

N. Koenig and A. Howard, et.al [14] is a well-known simulator that uses high-performance physics engines for rendering of indoor and outdoor environments. While Gazebo has rich features, it has limited abilities in creating visually rich environment of large scale and offer the realistic imagery. It has lagged behind various advancements in recent rendering techniques which allow photo-realistic rendering. A different class of approaches is based on game engines that enable rendering of photo-realistic camera streams

C. M.Itin, et.al [15] argues that experiential learning is the process of making meaning from direct experience, focuses on the learning process for the individual. It engages the learner at a more personal level by addressing the needs and wants of the individual. According to this definition, we design a story script for the learning service in Second Life.

## III. METHODOLOGY

The architecture of our simulator is shown in fig.1. It is composed of a multisensory model, controllers, scene datasets, a view synthesis module and a global visualizer. Our simulator is based on Robot Operating System (ROS) which has a modular design and can be customized, upgraded and reused. In the virtual environment, we first import the point cloud of the real scene, which is generated from a 3D reconstruction, into the ROS and show it together with camera poses of input images in Rviz, a 3D visualizer for the ROS framework. Then we control the virtual camera's movement throughout the virtual world and estimate its 6D pose by ROS in real time. The estimated pose is then taken as a reference to select the most similar color-and depth image pairs in a query input dataset. Next, we use the selected color-and-depth image pairs to synthesize the virtual view based on our view synthesis module. At the same time, the whole trajectory of the moving camera and synthesized colorand- depth image pairs are logged. In the following, we provide more details on the individual components of our simulator.

Our goal is to build a free-viewpoint photo-realistic environment for vision-based tasks. Unlike previous methods that build the whole virtual environment on perfectly reconstructed 3D geometry, our view synthesis module takes a sparse set of RGB-D images as the input and produces new color-and-depth image pairs from arbitrary viewpoints. It consists of novel depth refinement and view selection steps followed by a fast rendering process. These components work together to lower the rendering complexity and improve the quality of synthesized images.

Depth refinement: Pixel-accurate alignment of object boundaries between color-and-depth image pairs and accurate depth values are necessary for high-quality rendering. This is because inaccurate depth values and misalignment often lead to various visible artifacts, such as ghost contours. During offline preprocessing, we introduce a pixel-to-pixel multi-view depth refinement algorithm to achieve this goal.

ISSN PRINT 2319 1775 Online 2320 7876
Research paper  Volume 11, Iss 4, 2022

View selection: The quality of synthesized images depends not only on correcting misalignment between color-and-image pairs or filling holes, but also on selecting input views. Choose incorrect or redundant views based on angles or distances between two views, which often leads to blurring images. In order to avoid such cases, we select input images considering not only angles and distances but also overlaps between two views.

*Layered depth image based rendering:* We propose to use layered depth image based rendering to synthesize new colorand-depth image pairs. The core part of image based rendering is 3Dwarping which projects pixels in the reference image plane to theworld coordinate and then reprojects them to the newposition in another image plane using camera intrinsic and extrinsic matrices. However, when objects in the background and foreground are projected to the same position, objects in the foreground may be occluded by objects in the background, which is caused by incorrect depth information or reprojection errors. To solve this problem, we evenly divide the depth map into layers based on the maximum and minimum depth values. On each layer, we apply 3D warping with corresponding color-and-depth image pairs to produce new images and then use a median filter with a 3 × 3 window size to fill missing information in each new image. After that, we blend these new images together to produce the final synthesized image. Since layered depths have the ability to represent occluded elements, our approach better handles the visibility problem. We found four layers to be a good trade-off between quality and speed.

After blending, the synthesized image may be constantly subjected to holes that are caused by the fixed number of input views used. To address this issue, we propose an adaptive view selection approach using a variable number of input images to fill holes in the synthesized image. We first project a key image selected based on the angle, distance and overlap it has with the virtual view, to the virtual position, and then detect holes in the synthesized image. If the size of the largest hole is bigger than a threshold (e.g., 0.04% of the whole image), we then choose another input image to fill holes. We iteratively run this process until the size of the largest hole is smaller than the threshold or the number of input views reaches the maximum.

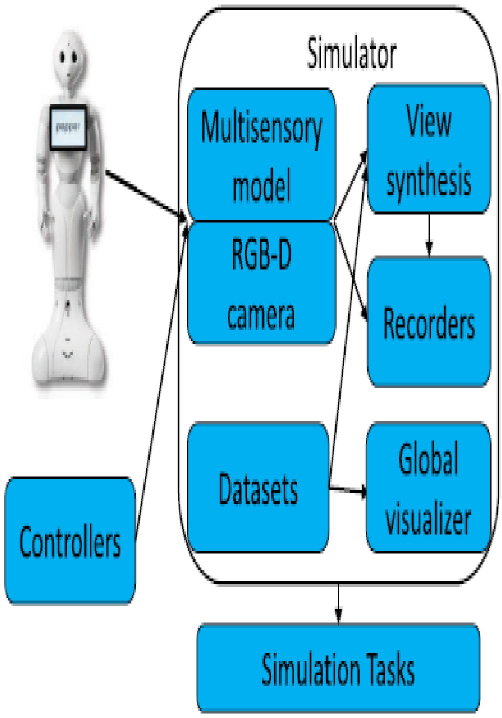


**Fig.1: The architecture of simulator.**

Multisensory Models and Controllers: PreSim is designed to investigate the issue of domain transfer from simulation to the real world. Thus, it is important for the multisensory model to be constantly subject to constraints of space and physics such as collision and gravity.

*Multisensory models:* We use Universal Robotic Description Formats (URDF) to describe multisensory models (e.g., humanoid robots). Therefore, the model and its properties can be configured (e.g.,

ISSN PRINT 2319 1775 Online 2320 7876
 Volume 11, Iss 4, 2022

types of sensors).As a demonstrator, we use the Pepper robot, which is a social humanoid robot from SoftBank.

*Integrated controllers:* We provide a set of practical controllers including joint state and navigation controllers to reduce the controlling complexity for the model's dynamic motions. The joint state controller is used to control the behaviors of joints of the model, including changing the pitch, roll and yaw angles. Our navigation controller allows controlling the model by directly sending movement commands. We also provide data recorders that allow saving all the data required by learningbased approaches. An example of a multisensory model and its trajectory.

## IV. RESULT ANALYSIS

We evaluate PreSim on seven static datasets including our three own datasets four datasets (Attic, Dorm, Playroom, Reading corner) from [4], and two dynamic datasets (Ballet and Break dancers). There are less than 220 color-and-depth image pairs in the seven static datasets that contain black and texture-less objects (e.g., white walls and writing boards), reflective objects (e.g., bottles and lights) and objects with small geometric details. Each of the two dynamic datasets contains dancing people with a sequence of 100 color-and-depth image pairs, captured by eight static cameras which are positioned along an arc at 20-degree intervals.

*Overall performance:* We randomly choose a color image from the initial captured dataset as our ground truth image and then use other images to synthesize the chosen image. Synthesizing one image (1280 × 720) takes 500–600 ms on a computer with 6-core Intel Core i7 8700 3.19Ghz CPU. Even though is faster than our approach which achieves 30 FPS, is based on a GPU while our method is free from the GPU.

**Table.1:Perforamnce Analysis**

| Parameters | AI | quasi-experimental |
|---|---|---|
| Time | 92.3 | 84.7 |
| Accuracy | 98.1 | 90.8 |

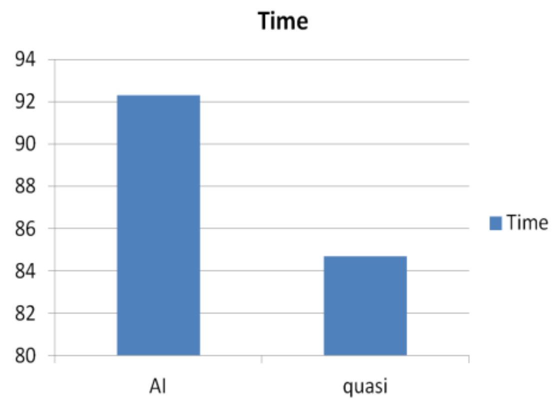


**Fig.2: Time comparison Graph.**

In Fig.2 time comparison graph is observed between AI and quasi-experimental.

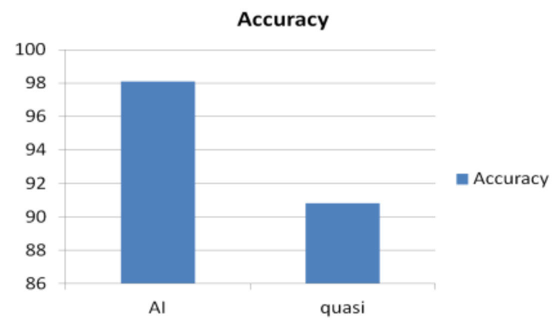


**Fig.3: Accuracy comparison Graph.**

In Fig.3 accuracy comparison graph is observed between AI and quasi-experimental.

## V. CONCLUSION

It is suggested that virtual environments and remote experimentation are suitable

ISSN PRINT 2319 1775 Online 2320 7876
*Research paper*  Volume 11, Iss 4, 2022

instruments for the collaborative process; they provide an intriguing viewpoint for teaching dispersed and collaborative across numerous applications. These are the kinds of technologies that have the power to greatly enhance immersion by fostering a sense of genuine presence and connection. The goal of the exhibition is to demonstrate how virtual environments and 3D remote experiments may be integrated to make science and technology jobs more appealing. The produced data may be utilized to train and evaluate data-driven methods for a variety of artificial intelligence applications, including 6D object pose estimation and depth estimation. Tests show that our simulator reduces the realism gap between the scene and the virtual world. Therefore, without requiring domain adaptation, vision-based algorithms created in the simulation may be used to actual physical platforms. Hence, by using this AI students will benefit in education like online class and information with in short period of time. Hence by using this AI better results are obtained interms of accuracy and reduction in time.

## VI. REFERENCES

[1] C. Wang *et al.*, "Densefusion: 6d object pose estimation by iterative dense fusion," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 3343–3352.
[2] M. Savva, "Habitat: A. platform for embodied ai research," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2019, pp. 9339–9347
[3] I. Laina, C. Rupprecht, V. Belagiannis, F. Tombari, and N. Navab, "Deeper depth prediction with fully convolutional residual networks," in *Proc. 4th Int. Conf. 3D Vis.*, 2016, pp. 239–248.
[4] P. Hedman, T. Ritschel, G. Drettakis, and G. Brostow, "Scalable inside-out image-based rendering," *ACM Trans. Gr.*, vol. 35, no. 6, pp. 1–11, 2016.
[5] M. Cavazza et al., "Causality and Virtual Reality Art," *Proc. 5th Conf. Creativity and Cognition*,ACM Press, 2005, pp. 4–12.
[6] J. Jacobson, "The CaveUT System: Immersive Entertainment Based on a Game Engine," *Proc. 2nd ACM SIGCHI Conf. Advances in Computer Entertainment Technology* (ACE 05), ACM Press, 2005, pp. 184–187.
[7] A. De Lucia, "Development and evaluation of a virtual campus on Second Life: The case of SecondDMI," *Computers & Education*, vol. 52, no. 1, 2009, pp. 220-233.
[8] D.C. Cliburn and J.L. Gross, "Second Life as a Medium for Lecturing in College Courses," *Proc. System Sciences, 2009. HICSS '09. 42nd Hawaii International Conference on*, 2009, pp. 1-8.
[9] C. Hong, "Service Design for 3D Virtual World Learning Applications," *Book Service Design for 3D Virtual World Learning Applications*, Series Service Design for 3D Virtual World Learning Applications, ed., Editor ed.^eds., IEEE Computer Society, 2008, pp.795-796
[10] P. Dev, "Virtual Worlds and Team Training," *Anesthesiology Clinics*, vol. 25, no. 2, 2007, pp. 321-336.
[11] S. Bronack, "Learning in the Zone: A social constructivist framework for distance education in a 3D virtual world," *Proc. Society for Information Technology & Teacher Education International Conference 2006*, AACE, 2006, pp. 268-275.
[12] L. Jarmon, "Virtual world teaching, experiential learning, and assessment: An interdisciplinary communication course in Second Life," *Computers & Education*, vol. 53, no. 1, 2009, pp. 169-182.
[13] R. Ortiz-Cayon, A. Djelouah, and G. Drettakis, "A bayesian approach for selective image-based rendering using superpixels," in *Proc. Int. Conf. 3D Vis.*, 2015, doi: 10.1109/3DV.2015.59.
[14] N. Koenig and A. Howard, "Design and use paradigms for gazebo, an open-source multi-robot simulator," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, vol. 3, 2004, pp. 2149–2154.
[15] C. M.Itin, "Reasserting the Philosophy of Experiential Education as a

ISSN PRINT 2319 1775 Online 2320 7876
 Volume 11, Iss 4, 2022

Vehicle for Change in the 21st Century," *The Journal of Experiential Education*, vol. 22, no. 2, 1999, pp. 91-98.