International Conference on Advanced Computing Technologies and Applications (ICACTA-2015)

3D Gesture-Recognition Based Animation Game

Niyati Gosalia\textsuperscript{a}, Priya Jain\textsuperscript{b}, Ishita Shah\textsuperscript{c}, Dr. Abhijit R. Joshi\textsuperscript{d,}\textsuperscript{*}, Neha Katre\textsuperscript{e}, Sameer Sahasrabudhe\textsuperscript{f}

\textsuperscript{a}Student, D.J.Sanghvi College of Engineering, Mumbai - 400 056, India
\textsuperscript{b}Student, D.J.Sanghvi College of Engineering, Mumbai - 400 056, India
\textsuperscript{c}Student, D.J.Sanghvi College of Engineering, Mumbai - 400 056, India
\textsuperscript{d}Vice Principal (Academics) & H.O.D. (I.T.), D.J.Sanghvi College of Engineering, Mumbai - 400 056, India
\textsuperscript{e}Asst. Professor (I.T. Department), D.J.Sanghvi College of Engineering, Mumbai - 400 056, India
\textsuperscript{f}Senior Faculty Associate, Tata Institute of Social Science, Mumbai - 400 088, India

Abstract

This abstract frames the research conducted on the existing version of a 3-D gesture-recognition based animation game, MathMazing with the purpose of overcoming the drawbacks of this system. In addition to this, we have summarized our approach to the proposed system.

MathMazing was implemented by our alumni [Sahasrabudhe et al, 2009] \textsuperscript{[1]} for teaching basic arithmetic to primary school. Our objective is to improve upon it. We have made a thorough study of its drawbacks and will be combating the same within the course of our project implementation. This endeavor not only requires the application of Logic and Kinesthetic skills but it also addresses spatial skills. This paper highlights the study of the existing system and puts forth a proposed solution that overcomes the limitations of the existing system.

* Dr. Abhijit R. Joshi. Tel: +91 9869353187
E-mail address: abhijitjoshi73@gmail.com

Keywords: Gesture based educational game; Blender game engine; Web camera; Exergaming; Multiple intelligences; Arithmetic drill.
1. Introduction

Electronic Learning (more commonly known as E-learning) refers to the utilization of electronic technologies to access educational curriculum outside traditional classroom. Typically, E-learning involves some kind of interactivity. Its major advantage is that it allows you to learn anytime, anywhere, provided that you have access to seem to be a chore for it makes learning a lively, interactive and more interesting experience, thereby enhancing the overall process.

The Department of Education stated, “Students who took all or part of their class online performed better, on average, than those taking the same course through traditional face-to-face instruction.” Students who mix online learning with traditional course work (i.e. blended learning) do even better (Internet Time Group Report, 2009). ‘Learners learn more using computer-based instruction than they do with conventional ways of teaching, as measured by higher post-treatment test scores.’ (‘Training and Retraining’ by Fletcher[2] and Kulik[3], 2000). Their study confirms that learners learn more using computer-based instruction than they do through traditional classroom methods.

The biggest advantage of E-learning is that it bridges the gap between the student and its instructor as none of them need to travel to each other, thereby eliminating expenses and inconvenience which is not the case in classroom learning. Teaching and communication techniques which create an interactive environment for learning and include demonstrations, games, animation, simulations on being compared to simple paper-pencil methods of study were found to be more beneficial. With E-Learning, students have more control over their learning process and can better understand the material, leading to a 60% faster learning curve, compared to instructor-led training. (Facts, Figures and Forces Behind e-Learning – August, 2000).

Statistics like the above are constantly proof to the fact that E-learning is increasing knowledge retention by a vast margin which is one of the primary benefits of this method of learning.

The collaboration of gaming and learning into a single application is done with the intent of making the user more interested in a particular subject, to motivate him to learn by constantly challenging himself/herself. Of course, this mode of learning could also have a negative impact if students tend to start showing signs of addiction.

Today, there is no dearth of educational computer games. But spending time on playing computer games leads to sedentary lifestyle habits, with associated health issues [Dietz & Gortmaker, 1993] [4]. Hence it is desirable to incorporate physical activity into computer games. Exergaming (exercise and gaming) is a new form of gaming that requires players to perform physical activities in order to play a game. Exergaming is in the form of bodily, physical movement and gestures and it comes into the scenario when the objective is to find alternatives to sedentary computer gaming.

Interactive gesture-based games used for educational purposes positively impact not only on children's learning, but also their enjoyment in physical activities. In addition, gesture-based games induce students to apply several abilities like logical, spatial and kinesthetic corresponding to multiple intelligences [Gardner H. 1983] [5]. Players need to solve the mathematical equations (logic), locate the path to the correct answer (spatial) and use gestures to steer the onscreen animated figure to the correct answer (kinesthetic).

Section 2 of this paper covers the study of the existing system along with its pitfalls. Section 3 presents the architecture of the proposed system. The paper ends with conclusion and provides some direction to future scope.

2. Study of Existing System

The system under study, MathMazing is a single player game having multiple-levels. It is a gesture-controlled game
consisting of a mathematical maze. It is used for teaching arithmetic to primary school students. It is developed using an open source 3D animation and game logic tool called Blender. The gestures are captured using Microsoft Kinect and imported into Blender through FAAST (Flexible Action and Articulated Skeleton Toolkit), a gesture-interpretation tool. FAAST is an open source middleware and is also compatible with many gesture recognition tools. FAAST helps in emulating keyboard input events which are triggered by gestures and player movement. It even facilitates the integration of the data captured by the input device, which is Kinect in this case along with the game logic module implemented in Blender.

The players of the game can view the instructions as to 'How to play' at the start. At each level, the player has to solve a randomly generated arithmetic equation, which will be displayed on the screen. Three choices are displayed at three different end-points of the maze, one of which is the correct answer. The player has to navigate through the maze to the correct answer, using gestures. The gestures decide the player movement through the maze. These gestures include some simple actions like bending forward/backward to make the 3D player character move forward/backward, or raising your left/right hand up to move the player left/right. Figure 1 shows the screenshot of the existing game in progress.

\[ \text{Fig.1: Screenshot showing the actual game in progress} \]

2.1. Workflow of MathMazing

The workflow of the existing system is depicted by Figure 2.

1. For the player movement, through the maze, we need to capture the player gestures. This is effectively done using a Microsoft Kinect.
2. The gestures captured by the Kinect are converted into signals for further processing and facilitating player movement through the maze.
3. A gesture interpretation tool, FAAST is used to interpret the signals captured by the Kinect.
4. The FAAST signals are converted to key press events.
5. The converted key press events are imported to Blender Game logic module for implementing it in the Blender engine.
6. According to the key press events imported to Blender game engine, the player movement is displayed on the screen.
7. The player can now decide the next step based on the current visual feedback received.
2.2. Pitfalls of the Existing System

1. Lack of Devnagari Support: The existing system lacks Devnagari support as the medium of game play, instructions as well as the numerals and equations are all in English. Due to this, students educated in Hindi medium schools, cannot follow this game.

2. Dependency on Kinect: Use of Kinect makes it highly inaccessible to the children/schools who cannot afford it as Kinects are expensive. Approximately a Kinect costs Rs12,000/- whereas a web-camera costs around Rs600/-. In our version we will be using a web camera instead of a Kinect. By doing this, the cost-effectiveness will increase.

3. Lack of Challenge: The existing game lacks in posing a challenge to players as it uses the same maze for all levels. This limits the difficulty level to only solving the equations quickly. Whereas if every level has a different maze, the player has to locate the path as well as solve the given equation within the given time-frame. A different sublevel for all basic mathematical operations (Addition, Subtraction, Multiplication, Division and BODMAS) will be incorporated to ensure thorough testing of student’s concepts.

4. Limited movements: The current version has very few movements (i.e., left, right, forward and backward). We will be adding more gestures such as jump to the game to increase its interactivity.

5. Bland user interface: The previously implemented user interface is not that appealing for the school-going kids. An effort will be taken to improve the interface so that, learning becomes Fun and Interesting.

3. Proposed System

We want to improve the existing version of MathMazing. This version will be implemented using a web camera in place of Kinect. The game environment will comprise a maze and a player. A basic mathematical equation (say, 1+2+3=? ) will be displayed on the screen. At three different checkpoints of the maze, one distinct option will be displayed per checkpoint (say, 5, 7 and 6 which is the correct answer). The objective of the game will be to reach the correct answer by moving around the maze using gestures. The game will have a different level for every mathematical operation (Addition, Subtraction, Multiplication, Division and BODMAS). Every level will have a different maze unlike the existing version which will make the game more challenging. Also, the game will be scripted in Devnagari to increase its reachability.

3.1. Implementation Plan

We plan to use an open source 3D Animation tool called Blender for designing the game environment. Being a gesture-based game, the gestures of the player will have to be captured and processed in order to detect the gestures. Web-camera will be used to capture the gestures of the player. OpenCv software will be used for processing the captured gestures in order to recognize them. OpenCv (Open Source Computer Vision) is a library of programming functions mainly aimed at real-time computer vision. OpenCv will be integrated with Blender in order to generate
the key press events that will facilitate player movement in the game.

3.2. Architecture of Proposed System

The proposed system's architectural design was made keeping in mind the primary goal of developing a game with an effective gesture recognition and detection system for human computer interaction. In order to achieve this goal, two main tasks will be carried out. The first task involves determining the position and orientation of the hand in 3D space, in each frame. Secondly, the hand and its gesture will be recognized and classified to provide the interface with information on the necessary, corresponding actions that will be required to be taken in response the gesture the hand is making. Thus, the hand must be tracked to give positioning information to the interface and the gestures must be recognized to provide the necessary feedback to the movements as well as hand gestures. The interface is responsible for giving this kind of a response to the player and his/her actions.

The architecture of our system will be as shown in the Figure 3.

i) Background Subtraction: The implemented system architecture starts with the background subtraction. It is a technique in the fields of image processing and computer vision wherein an images foreground is extracted for further processing. For example: Object Recognition.

ii) Image Segmentation: After the subtraction of the captured image and detection of hand, the segmentation, tracking and recognition are done. Segmentation is the process of partitioning a digital image into multiple segments.

iii) Gesture Recognition: The image that has been processed will be compared with a repository of the pre-recorded gestures and their corresponding key-press events. This facilitates player movement on the maze.

5. Conclusion:

In this paper we have presented a study of a 3D gesture-recognition based E-learning game, called MathMazing, used for teaching arithmetic to primary school students. E-learning games make studies effective, practical and also lead to better retention. "I hear and I forget. I see and I remember. I do and I understand."(Confucius). Our objective is to overcome the drawbacks of this system. We have stated the working of the existed system and its pitfalls. Also,
we have put forth our proposed system which will overcome the drawbacks of the previously implemented version.

**Acknowledgements:**

We wish to express our gratitude to our Project Guide, Dr. Abhiit R. Joshi for his overwhelming encouragement and helpful feedback. Also, we would like to thank Prof. Sameer Sahasrabudhe & Prof. Sridhar Iyer who gave us the opportunity to do this research project in Blender under the Indian Institute Of Technology. We are grateful to our mentor, Prof. Nitin Ayer and our co-guide, Prof. Neha Katre for their patient guidance. Elsevier has provided us with a podium to publish our paper, and we are sincerely appreciative of the same. Lastly, we would like to thank the authors of various reference materials mentioned in the references for their commendable research.

**References:**

1. S. Sahasrabudhe, A. Shah, M. Thakkar, V. Thakkar & S. Iyer, (2012). MathMazing: 3D gesture recognition exergame for arithmetic skills. In Proceedings of 20th International Conference on Computers in Education (ICCE) 2012 (pp. 11-17). Singapore: APSCE
2. Fletcher, J. D., & Tobias, S. (2000) (Eds.). Training and retraining: Handbook of Research on Educational Communications and Technology. A handbook for business, industry, government, and the military. New York: Macmillan
3. Kulik, J.A. (1994). The Impact of Education Technology on Student Achievement. Meta-analytic studies of findings on computer-based instruction. In E.L. Baker, and H.F. O’Neil, Jr. (Eds.). Technology assessment in education and training. Hillsdale, NJ: Lawrence Erlbaum
4. Dietz, W., & Gottmaker, S. (1993). TV or not TV. Fat is the question. Paediatrics, 91, 499-501
5. Gardner, H. (1983). Frames of mind: the theory of Multiple Intelligences. New York, NY: Basic Books.
6. Alharthi R., Karime A, Al-Osman. H. and El Saddik. A, (2012). Exerlearn bike: an exergaming system for children’s educational and physical well-being, IEEE International Conference on Multimedia and Expo Workshops, 9th – 13th July 2012, Melbourne, Australia