Interactive Augmented Reality-based System for Traditional Educational Media using Marker-derived Contextual Overlays

Grishma Alshi, Mansi Dandiwala, Mikhail Cazi, Renuka Pawar
grishma.alshi@gmail.com, mansi.dandiwala@gmail.com, mikhailcazi@gmail.com, renuka_pawar@spit.ac.in
Sardar Patel Institute of Technology,
Andheri, Mumbai

Abstract — **Augmented reality is making an impact across different sectors of society. Drawing from this fact, the system proposed herein aims to aid educators by making the process of learning interactive, enjoyable and effective. Using this, theoretical contents of books are enhanced using computer vision, 3D models and Human-Computer interaction (HCI). This enhancement is made available to students in the form of a mobile phone application through an AR interface. One can hover over data or information in the book and get the augmented experience in terms of 3D models, audio or videos. Augmented reality is fast becoming an important part of human life with emerging technologies such as tracking techniques, computer vision systems, graphics, mobile computing gaining traction. The proposed approach aims to interesting and motivational compared to old teaching methods. The paper also discusses the educational settings where the suggested model can be used, along with a detailed description of the prototype, its applications, and responses.**

Keywords—Augmented Reality; Interactive Learning; Markers; Tradition Teaching Practices.; Overlay; 3D object

I. Introduction

AR revolves around the idea of superimposing various elements over the user's view as seen through a device camera to form a coherent and communicative experience. During its embryonic stage, simulators[1] presented visuals, sound, smell and/or vibrations, but using bulky devices or otherwise complicated setup. However, in the recent times, AR is made available to the masses over their smartphones with a better graphical overlay. Due to this ease of access, a notable rise in the use of AR for various fields can be observed since the past couple of years[2]. Along with traditional teaching methods, interactive learning has also begun to be employed. A survey carried out reflects that 84.8% of the participants support the idea of including AR into teaching practices. Interactive learning involves urban computing and personalized digital media to enhance the user's learning experience. Its major advantages are elasticity (expanded information) and adaptability (user-centric information provisioning). These advantages make it possible for AR to be incorporated into interactive learning. Allowing a user to access highly graphical models and other associated data will strengthen their understanding of the matter read by them from a book.

This paper is centered around the idea of incorporating AR into the everyday educational practices, entailing an interactive learning approach. This system can utilize commonly available Android-based[1] smartphone camera to capture and detect patterns called as markers. These markers are predefined, specific to the book curated for the project. Various charts, image and text boxes are used as markers that work as initiators for the augmenting overlays of related data. These overlays suggest related information in the form of audio, video, text and internet search results are still made available to the user while using the smartphone camera. In order to visualize new concepts, 3D models built will be used to make the application a lot more interactive and eye pleasing.

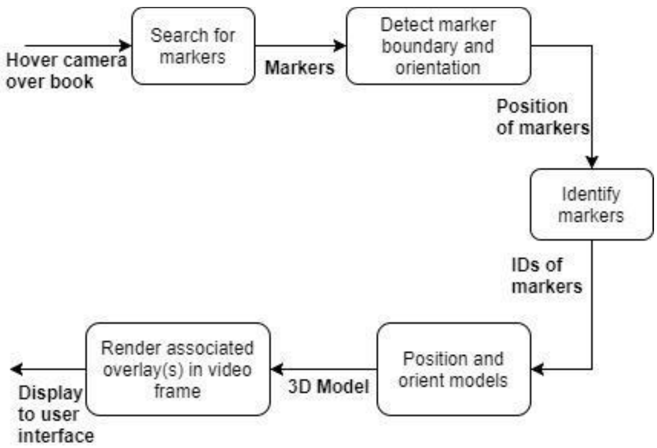

Fig. 1: System Overview

II. System Definition

A. Augmented Book

Augmented reality books appear to be just like any ordinary book, however, when a mobile phone[7][11] running given system has hovered over it, 3D models, audio and, /or video is made available to the user in the form of an overlay. It can also comprise of other interactive components, such as haptic feedback. These overlays help the users to understand various topics in a more interesting and efficient manner. Such books also help students to visualize data in detail and improve their visual-spatial ability. They also contain markers which serve as links to the overlay content. While adding a book to the system database, different images or characters may be selected for marker detection[12]. These images are stored in the database for the purpose of template matching.

A. Markers

Marker[4] in reference to augmented reality is any object or image located somewhere in the scene to refer to a fixed point of a position in the real world. These markers act as visual hints in for the augmented content to be displayed. A good marker should be such that any system can actively detect as well as distinctly identify it. Markers help to determine the pose of the camera to be in the real scene. The pose of the camera depends on both the position and orientation of the marker. The method of calculating this pose based on the position and orientation of the marker is called tracking. Various types of markers are:

1. Marker with black boundaries - These are square or rectangular patches with a thick black border around it. Recognizing these rectangles in pictures or images has been refined over the years. While tracking these markers, the system checks for black borders around the image. For the border detection, thresholding is used. The marker patterns are stored in a centralized database. The encapsulated part within the border is compared with the stored templates from the database, for verifying the real-world markers. Based on the marker detected, the relevant augmented content is displayed.

2. NFT marker - Natural feature tracking markers are also images, without any significant black borders. Instead of detecting the black border, the natural features of the source image are directly processed. However, defining these features and accurate extraction of the same is quite difficult to do. The features extracted are stored in a centralized database for template matching. This, these markers are complicated to work with when compared to the former type.

3. GPS based marker - GPS based markers are not solely dependent on the user's current location but it can also mark the location of some building or any such object using GPS. Relative to user's position, the augmented objects can be displayed up to few meters. The accuracy of these markers is low. For better results, GPS markers are usually combined with other methods of tracking.

Tracking can also be achieved using markerless approach. Markerless augmented reality can use any of the real world objects as a target based on which the augmented content can be displayed. Markerless systems have the ability to extract and capture information from their environment Markerless AR makes use of GPS feature on the mobile phone to find the location of the objects.

This system is intended to complement the existing tradition approach i.e. reading from books, and hence marker-based tracking as the images, graphs, and illustrations in the sourcebook could be used for precise tracking.

Fig. 2: Markers with black boundaries

B. Types of visuals

The augmented content displayed through the application on the mobile phone can be audios, videos, 3D models[9], as well as web links related to the topic, along with a manual web search option. Audio here indicates a set of some recordings that will play in the background. In case of videos, a thumbnail for the same will be displayed in a pop-up on the screen, which the user can click on it to watch the videos. Other visual elements include 3D models, created using Unity3D and Blender3D. Standard visual elements include hyperlinks to the web pages related to the topic or image depending on the marker.

III. IMPLEMENTATION

The process flow for initializing and implementing the system begins with content creation. This book contains the markers for selected words, characters, images or charts. Once a book's marker has been added to the centralized database, users may start using the system by scanning the marker[3] via the smartphone application.

After the image is captured, marker detection is done to determine what augmented content is to be displayed. The algorithm given below is used to detect the markers:

1. Divide the scene containing your marker into regions. Each region to be divided into sub-regions, which are of all same size.
2. Track the edges in all the regions.
   All edges in these regions are identified using Gaussian derivatives. The Gaussian derivative helps in convolving the image into both horizontal and vertical directions with integer-valued separate filters. The threshold for convolution is predefined. All the edges in the image showing strong maxima i.e. greater than the threshold are considered for marker detection.

After the Gaussian derivative is applied, Sobel operator is used on the detected edges, which helps in determining the orientation of the edge.

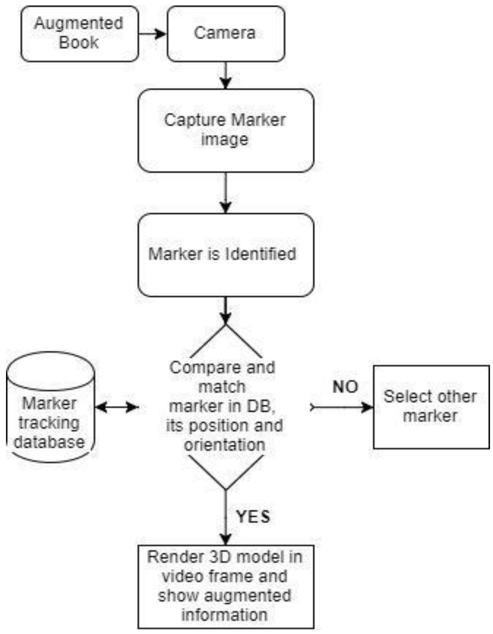

Fig. 3: System Flow Diagram

3. Determine segments from the detected edges.

RANdom SAmple Consensus-grouper (RANSAC) algorithm is implemented on these edges which leads to the creation of line segments in the given image containing the marker. Inliers of the marker are most important and to be found using this algorithm. RANSAC Algorithm:

i. Consider two points at random which lie within the same region, making sure that the orientation of these points is consistent with each other to complete a line.
ii. For any line, the point can become a part of it only if its orientation is consistent with that of the line and close to the same.
iii. Previous steps are repeated until we find distinct segments.
iv. Lines with enough number of points on it are considered as the segments.

4. Join the segments found in the previous step. Two segments can be joined only if they lie near to each other, their orientation is consistent and if the pixels between the two segments are on an edge, which is found using the same criterion as used in step 2.
5. Elongate the lines in the scene, forming merging segments along the edges found from the previous step. Check if every pixel on the prolonged line is on an edge using the previous edge detection technique. When the observed pixel is not present on edge, move further away to a pixel of a White colour than that of the edge is found which can be the corner of the marker.
6. Keep only the lines with corners.
7. Find the makers by making use of lines and corners. For this, find a chain of 3-4 lines. A chain is a collection of lines in which the end of one line touches the other line at its start. Now only keep the chains which form a rectangle.

When done with detecting all the markers, 4 intersections of the lines are acknowledged as corners for each of marker. A robust algorithm for marker detection, which guarantees reliable detection and identification is obtained. Because of this, even if only 3 of the 4 lines gets detected by the camera, the algorithm is capable of detecting the marker and camera pose without any miscue.

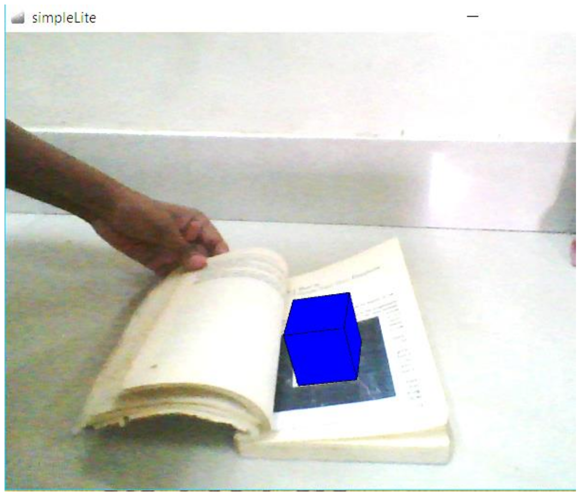

Fig. 4: Display cube using black boundary marker

After the marker is identified, the system will match it with the existing markers or templates in the database. If the match for the marker is not found user may capture the image of another marker, or select some other media entirely. If a match is found, the system will render the relevant augmented content.

IV. SYSTEM COMPONENTS

A. Unity3D

Unity is a type of software called a ‘game engine’. Although primarily used to program, develop and create simulations and video games (both 2D and 3D), Unity[7] can be very helpful in creating AR features and applications. It provides the ‘environment’ in which AR applications can be built. The scene of a Unity application acts as the environment in which markers and their corresponding objects to be displayed on the screen are set up and linked. This linking is done with the help of other programs and software like Vuforia.

The below picture is a scene in the Unity game engine, which is translated into an AR scene when it is run. The images of the grass and concrete pattern are the markers in the scene. The camera on recognizing these markers will display the objects above them - in this case, a spaceship.

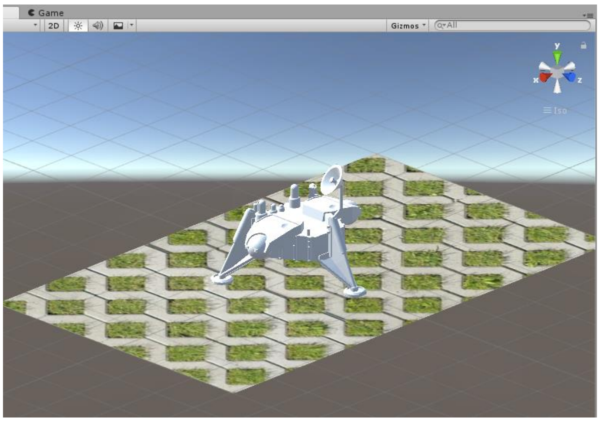

Fig. 5: Spaceship augmented over pattern

B. *Vuforia*

Vuforia is an SDK (software development kit) for AR that enables the creation of AR applications. Vuforia provides the functionality of recognizing and tracking targets or markers in the real world, by using the technology of computer vision. The targets supported are a variety of 2D and 3D types including markerless images, specific markers, and simple 3D object configurations. It then allows for the positioning and orienting of virtual objects into the real-world environments. Vuforia[6] also acts as the database or dictionary which links the targets/markers to the corresponding objects to be displayed in the AR application.

C. *Blender3D*

Blender is a computer graphics software which can be used to create models and art for AR applications. In the process of making an AR application, Blender[8] is used to create the virtual objects which are superimposed on the real world. Blender allows the modeling[10], texturing, and animation of 2D and 3D objects.

D. *MyMultimediaWorld*

MyMultimediaWorld is an online asset management system. It offers complete multimedia content diffusion services, which includes the support of images, videos, audios, 3D objects, and games, as well as their combination and linking into complex scenes. The platform also offers APIs to integrate extraction algorithms using resources uploaded onto the service. Thus, MyMultimediaWorld is a useful application for linking our target markers and objects to be displayed. The service acts as a database or dictionary for our AR application. The code behind the linking in MyMultimediaWorld can be easily changed and functionalities updated.

E. *ARAF Browser*

ARAF stands for AR Application Format. An ARAF Browser is capable of reading and decoding AR content like images, videos, and objects into an application. The ARAF Browser uses sensor information from a device's GPS[5], camera, accelerometer, etc. and thus connects the real and virtual world. It provides for commanding actuators and implements dynamic content with the use of interpolators and scripting. A user can interact with the application using the device's touchscreen. In our project, the ARAF browser is used to view the real world with the device camera. The Browser searches and recognizes image targets which have been stored in the MyMultimediaWorld 3+base. On recognition, the browser follows the code given in the MyMultimediaWorld object - and does a follow-up action, usually playing a sound, or overlaying a virtual object.

V. USER SURVEY

In order to understand whether the current education system was in need of an interactive AR based teaching method, a survey was carried out by us. This survey was carried out by distributing a link to the participants. A total of 4 questions were asked pertaining to study methods, AR knowledge and its acceptance as an aid to traditional teaching methodology.

The survey received responses from 66 individuals with the following respondent types: teacher, student or professionals. The observations gathered from the responses were crucial to us as this data helped in while creating the book for our application. The first question asked the respondent to name the teaching method that they prefer the most. It wasn't astonishing to see 55.5% respondents to name visual/interactive methods as their preference over traditional media. Supporting this view, 95.5% respondents believe that visual aids help students understand a concept better. While being questioned specifically about our subject, 83.3% respondents do know about the emerging technology of AR and 84.8% are open to the infusion of AR in teaching practices. The complete set of survey questions are as follows:

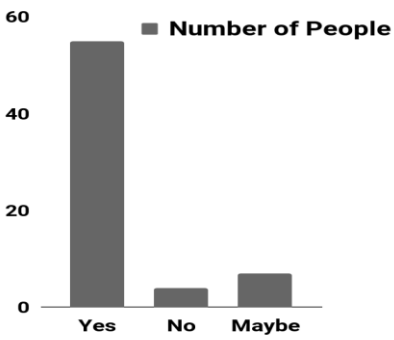

Fig. 6: Do you know what augmented reality is?

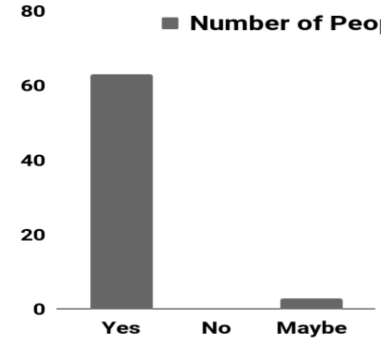

Fig. 7: Does visual aid help students in understanding a concept better?

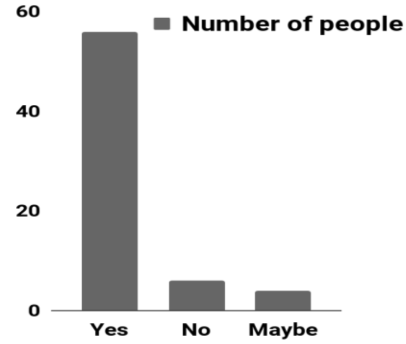

Fig. 8: Would you be open to AR infused material for educational purposes?

VI. LIMITATIONS

Currently, the system, being a rudimentary implementation of the idea proposed in this paper, has certain limitations. The proposed model is for Android-based operating smartphones making the screen size limited to the smartphone model. Also, for the system to work a working internet connection is needed. Good internet connectivity ensures no delay in overlaying the augmenting layers over the screen. Another point to consider is that this application requires students to make use of devices to access information, creating a dependency on digital devices.

VII. CONCLUSION

This paper presents an outline for an educational application based on the technology of AR. It was established that innovative methods like visual-based learning and interactive learning make a huge difference in the comprehension and retention of knowledge in students. Such teaching methods also increase student participation in classes. Also, the learning process is quickened along with increased intellectual curiosity of students.

The application converts a banal textbook into an informative, interactive 3D object - increasing the ease of comprehending complex concepts. This method removes the problems associated with language barriers. Also, learning materials are easily available without too much expense. The various methods of locating and tracking the markers, and the associated core system design was prototyped, along with the architecture and the various software components used to build it.

VIII. FUTURE WORK

Although using this technology, the transition from real world to augmented world is easy and very helpful, some of the users may find it cumbersome because of limitation of Android mobile device support alone. In the future, to broaden its reach, the system will be extended to other devices[5] as well. Also, the quality of information and accuracy of marker detection will be improved. The complexity of graphics is intended to be reduced, as well as making the augmented content more realistic. Multi-marker tracking method will be implemented in future to further improve the AR tracking, so that user may look for one or more than one marker at the same time.

REFERENCES

[1] Julie Carmigniani, Borko Furht, Marco Anisetti, Paolo Ceravolo, Ernesto Damiani and, Misa Ivkovic; *“ Augmented Reality technologies, systems, and applications”;* Springer 2010

[2] J. Michael Spector, M. David Merrill, Jan Elen and, M.J. Bishop; *”Handbook of Research on Education Communications and Technology”*; Springer 2014

[3] Yan Fei Gao, Heng-you Wang and, Xiao-Ning Bian; *“Marker tracking for video-based augmented reality”*; IEEE 2016

[4] Sanni Siltanen; *“ Theory and applications of marker-based augmented reality”*; VTT 2012

[5] Dieter Schmalstieg, Tobiad Langlotz, and Mark Billinghurt; *“ Augmented Reality 2.0”*; Springer 2011

[6]Fuguo Peng, and Jing Zhai; *“ A Mobile Augmented Reality System for Exhibition Hall Based on Vuforia”*; IEEE 2017

[7]Sung Lae Kim, Hae Jung Suk, Jeong Hwa Kang, Jun Mo Jung, Teemu H. Laine and, Joonas Westlin; *“ Using Unity 3D to Facilitate Mobile Augmented Reality Game Development”;* IEEE 2014

[8] Pooja Bhawar, Nitin Ayer and, Sameer Sahasrabudhe; *“ Methodology to create optimized 3D model using Blender for Android devices”;* IEEE 2013

[9] Adrian Clark and, Andreas Dünser; *“An Interactive Augmented Reality Coloring Book”*; IEEE Symposium 2012

[10] Jeff K.T. Tang, Tin-Yung Au Duong, Yui-Wang Ng and, Hoi-Kit Luk; *"Learning to Create 3D Models via an Augmented Reality Smartphone Surface"*; IEEE 2015

[11] Qingfeng Zhang, Weilong Chu, Changhong Ji, Chengyuan Ke and, Yamei Li; *"An Implementation of Generic Augmented Reality in Mobile Devices"*; IEEE 2014

[12] Jonathan J. Hull, Berna Erol, Jamey Graham, Qifa Ke, Hidenobu Kishi, Jorge Moraleda, Daniel G. Van Olst, *"Paper-Based Augmented Reality"*, IEEE 2007