Augmented Reality with 3D Human Face

Archana B. Patankar
Asst. Prof., Computer Engineering Department,
Thadomal Shahani Engineering College, Mumbai

Onkar A. S.
Lecturer, Computer Engineering Department
Thadomal Shahani Engineering College, Mumbai

ABSTRACT
In this paper, the design of intelligent agent is discussed. This agent is used to percept the changes in the environment and changes the expression of 3D human face using 3D morphing. The design is proposed for implementation in Java 3D API. Type of intelligent agent is a simple reflex agent and it has four functions, the first is to percept the change in the environment, the second is to understand the change and the third is to perform actions by generating expression on 3D human face as actuator.

General Terms
3D Facial Animation, 3D Objects Morphing, Java 3D API, Virtual World, Real World, Color Detection, Facial Expression.

Keywords
Augmented Reality, 3D Facial Animation, Intelligent Agents, Java 3D API, Image Processing, Artificial Intelligence, Intelligent Systems, Computer Vision, Human Computer Interaction.

1. INTRODUCTION
The augmented reality is an area in which one should create a bridge from the real world to virtual world. In this type of reality, the change in the real world should escort to change in the virtual world provided that the changes will be guess or identify by the system. The proposed agent is a type of simple reflex agent. In this category of agent the environment is percept or sense by using sensor. The sensed information is further used to do necessary actions by the actuator.

In 3D animation industry, the major interest of research needs to carry out the realistic expression of the 3D character model face. The 3D facial animators are appointed to work on the task of getting correct expressions of face. Also, it requires a thorough understanding of facial details and emotions. For an instance of time, humans may produce few unique facial expressions with fewer efforts. To mimic those expressions in the virtual world is really tough and major dispute. Also, to become a facial animator, the person should practice to express emotions and some study to react in a particular situation.

Now days, the useful research tools and freeware are available over the Internet, who helps user to perform the work quickly in short development time. FaceGen is software downloaded from Internet to generate 3D face objects with major facial expressions and phonemes. In the proposed system, all the 3D human facial expressions are created with the help of FaceGen software. FaceGen allows building 3D virtual face by providing profile images of any person (Front, Side profile images. These images are selected to generate 3D human face with all expressions like anger, surprise, sad etc. In order to generate 3D object files for every specific expression, one should make changes or modify those objects by exporting them to Adobe Maya [1] or 3D Max software.

Every color is associated with some expression. Consider an example, when one say “red color”, it is color indicating anger, similarly when one say “green color”, it is color indicating greedy expressions etc. There are so many colors and each of them will have specific understanding and relate to specific color. It will again differ from person to person. In this paper the color presence is identified through processing of the captured real time images and then associates the identified color with its related facial expression by an intelligent agent.

2. BACKGROUND
A facial animation is a basic area of computer graphics which encapsulates models and techniques that are used for generating and animating the images of human head or face. It includes variety of techniques from morphing to 3D modeling and rendering. In 2D, the transformation of one still image to another still image has transitional images are generated in between the provided pair of images. This technique is called image morphing and it was used in late 80s for performing facial animation in movies. A morphing consists of geometrical deformation which aligns the target image and cross fade technique is used to create smooth transition effect in the texture of image. For every intermediate transitional image as a key frame, an acoustic node is added in order to create a speech based animation.

Another method of a facial animation is in which the sequence of images captured and concatenated from the video [2]. This is used in video rewrite uses computer vision techniques to track the lip movements from the video where existing footage of an actor is cut into the segments corresponding to the phonetic units which are blended together to create new animations of a speaker. A 2D animation techniques used only supports an animation of lower part of human face. Further these are composited with the video of the original actor to produce the final animation.

A 3D facial model has powerful set of methods that would bring into being realistic animated expressions and visual appearance [3]. The traditional 3D animation methods used shape or morph target, skeleton muscle systems, bones or cage methods, motion capture using points on the face, and knowledge based solver deformations.

A Morphing is a name coined over a decade ago, for animation sequences which display gradual transformation. We refer to the 3D bodies which are represented polyhedral objects, i.e. by vertices, straight edges connecting them, and faces, each consisting of a planar vertex set with the connecting edges. This is the most common representation for 3D bodies in computer graphics, and especially in modelers for 3D animation. A 3D morphing generates the intermediate 3D objects and rendered them in a sequence to get continuous animation image sequence to depict the transformation.

3. DESCRIPTION

The hearing-impaired people can get benefit from the interactive talking human face by means of visual speech. Development of the lip reading tutors for deaf people and the language tutors for new learners & children need the visual and voice synchronization and synthesis. Now days, the user interactivity has come into the systems to make interesting human machine interface [4]. The animated characters will not only imitating humans but also look and behave like them. These facts drive the research into emotional and an intelligent virtual character which are expert enough to think and act rationally to react on the provided change in its environment.

The applications provided above have scope for face synthesis. Also have a great prospective towards interactive user interfaces deployed later into the e-learning based systems. These character talking heads can be used for low bandwidth teleconferences and virtual meeting rooms instead of the actual users and work as animated representatives [5]. The character animation and talking face has a major scope in movies, advertising, television commercials and gaming industry. The goal of this thesis is to animate the realistic 3D human face which will realistically produces the facial expressions. After perception from the environment, the character face produces the resultant expression, which will interact with the user in visual form. Another challenge will color recognition with real time camera capturing.

3.1 Design of Intelligent Agent

The design of a 3D human face is the important part of this system. A 3D human face is modeled in Maya software. In the modeling of the character face, the set of expressions will be modeled by taking replicas of the original 3D face and making slight modifications nearby eyes and lips. Every modeled expression, phoneme and original 3D objects will be exported to object files.

A simple reflex agent as an intelligent agent (see Figure 1) will be implemented has the following components:

Input: Set of different color cards and hand signs, images

Sensor: Camera

Actuator: 3D human face

Output: Instructions, expressions & audio-visual outputs

Now days, the creation of the human face need to understand the various components of face called as facial features [2]. Most of the 3D animation software tools used to model the 3D face and control points are marked by taking reference of generic face models from the available library.

3.2 Detailed Scheme of Implementation

The proposed system has the following stages:

1) Images will be capture and store using real time camera which is used to sense the environment.
2) Color recognition will later perform on any random image from the stored images.
3) To define & develop the set of instructions and set of expressions for an intelligent agent, it will match the color and its corresponding expression. For this the instruction command like “anger”, “sad”, “surprise” etc. will be add to instructions text file.
4) Modeling and development of 3D human face with necessary objects files generate by 3D modeling software like Maya etc.
5) Morphing using Java 3D API, in this step, one object file to morph to another object file. This morphing output will be display in the animated form to mimic the realistic expression with facial animation.

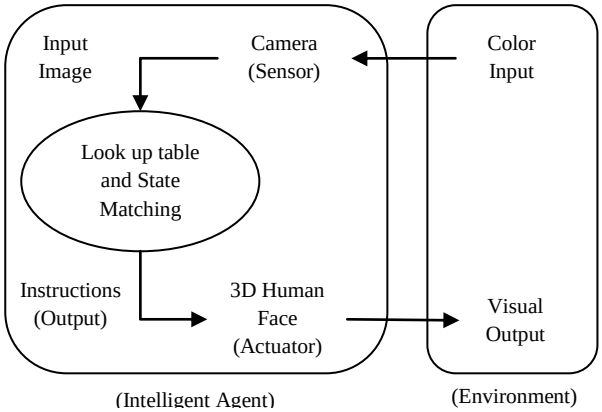

Fig 1: Design of Intelligent Agent

3.3 System Requirements

Software: Java 3D API, FaceGen, Java NetBeans, Maya, Matlab (experiments)

Hardware: Camera, computer with advanced graphics compatibility.

4. IMPLEMENTATION

For implementation of this simple reflex intelligent agent, the following steps are involved shown in block diagram:

4.1 Real Time Camera Interface:

In order to percept the changes in the real world, the real time camera interface is provided as discussed in the following block diagram:

Fig 2: Real Time Camera Interface

4.2 Color Recognition:

The following block diagram explains the working of color card recognition:

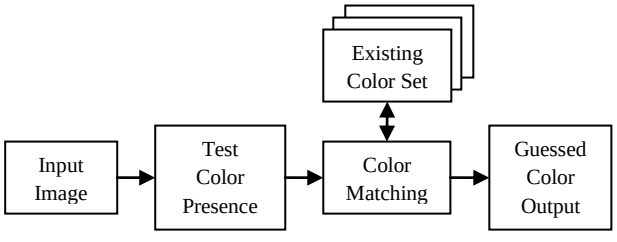

Fig 3: Color Card Recognition

It uses any random image from the directory of output images and it will prepare red, green, blue planes from those planes the maximum intensity level is calculated from each plane. Further those values are used as red, green and blue values of color presence. Finally, based on maximum red, green, and blue values the color is decided. The output color name is stored in file that will be used as instruction to generate expression.

Here, the color is identified which is also displayed in window to show the color guessed from the environment. To calculate maximum red, green and blue values, total numbers of pixels are counted for individual intensity in red plane or green plane or blue plane and maximum value is considered as maximum value of red or green or blue color. These are the three values which is used to denote the color in RGB form.

From RGB values, the corresponding expression command will be added to instructions file which is used by expression generator.

4.3 Expression Generator:

In this the morphing is performed in which source 3d object file is morphed into another target 3d object file. It uses the command specified in instructions file and output is produced in java 3d applet.

The above block diagram shows the 3d morphing to generate open smile of 3d face and back to the previous straight face expression.

5. RESULTS AND DISCUSSION

The camera is used to capture the real world images .The result of image capturing program written in java is shown below:

Fig 5: Camera Capturing

The captured images are shown below (50 images are captured):

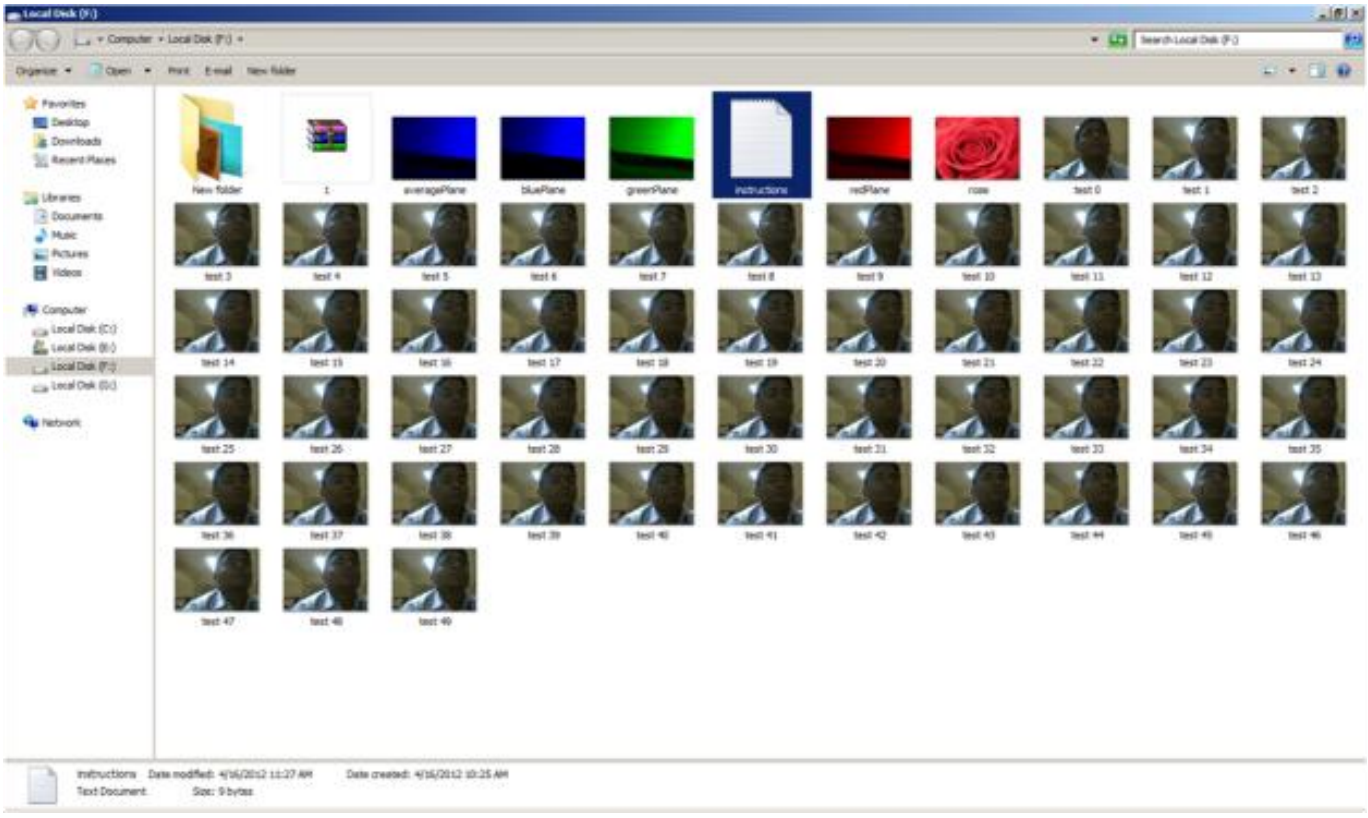

Fig 6: First 50 Images Captured

After some time, the images will overwrite by the new 50 images as shown:

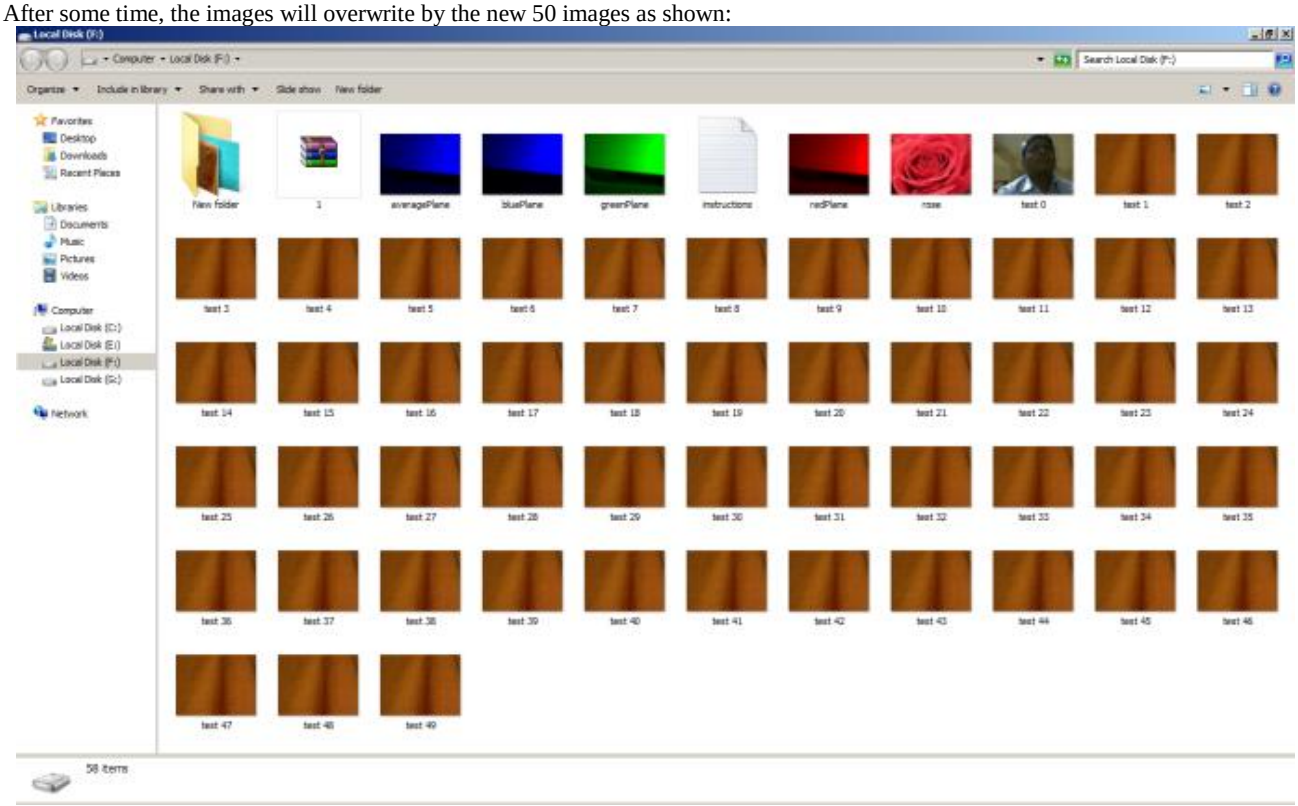

Fig 7: Images Overwrite by next 50 images

5.1 Result of Color Detection:

Randomly one of the captured images will select which is shown below:

Fig 8: Selected Random Images among 50 images

From the above image color will guessed as follows:

The guessed color has RGB values as red=153, green=73 and blue=0. The corresponding color is shown in the window panel in Fig. 9.

Fig 9: Guess Color by Color Recognition Program

Depending on the color produced if the red color is in a range of 120 to 255 then "anger" is the instruction will be added to the "instructions.txt" file as shown below:

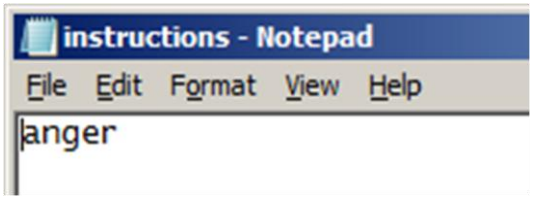

Fig 10: Instructions written in text file

To instruction is to generate the morphing output, the result of morphing program written in Java 3D using 3D transformation is shown below:

5.2 Results of 3D objects Morphing using 3D Transformation:

Initial Expression: The following is the output screen displaying initial expression:

Fig 11: Initial Expression (Straight)

Target Expression: The following is the output screen displaying target expression:

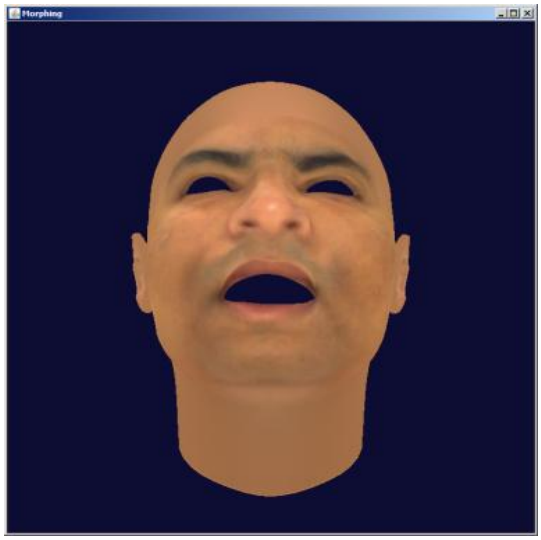

Fig 12: Target Expression (Anger)

In the output window, the initial expression is morphed into the target expression in animated form. In this test version of the project the eyeball, teeth and tongue is not produce the morphing output.

6. CONLCUSION

Thus, the real world is mapped to virtual world in order to achieve augmented reality. A proposed agent is a kind of simple reflex agent. The color card is put in front of the camera mounted nearby screen. The camera senses the environment and will produce the image using image capturing applet. This image is used by color recognition program. This unit checks for the provided input image and will match with its corresponding instruction based on the guessed color. This instruction will be provide to the 3D human face which will generate the expression output.

The project will be further extending to produce output language of an agent as any Indian regional language like Marathi, Hindi etc. The agent will be used in augmented reality applications to identify a person or an object in real time environment comes in the front of the camera.

7. REFERENCES

[1] António Moura, Ingrida Mažonavi, João Nunes, Justinas Grigaraviius, "Human lips synchronisation in Autodesk Maya", Systems, Signals and Image Processing, 2007 and 6th EURASIP Conference focused on Speech and Image Processing, Multimedia Communications and Services, 14th International Workshop, Page No. 365-368, 2007.

[2] G. N. Kodandaramaiah, M. B. Manjunatha, S. A. KJilani, M. N. Giriprasad, R. B. Kulkarni, M. Mukunda Rao, "Use of lip synchronization by hearing impaired using digital image processing for enhanced perception of speech", Computer, Control and Communication, (IC4 2009), 2nd International Conference, Page No. 1-7, 2009.

[3] Kaihui Mu, Jianhua Tao, Jianfeng che, Minghao Yang, "Real-Time Speech-Driven Lip Synchronization", IUCS, 2010.

[4] S. Lokesh, G. Balakrishnan, S. Malathy, K. Murugan, "Computer Interaction to Human through Photorealistic Facial Model for Inter-process Communication", International Conference on Computing Communication and Networking Technologies (ICCCNT), Page No. 1-7, 2010.

[5] Kyung-Geune Oh, Chan-Yul Jung, Yong-Gyu Lee, and Seung-Jong Kim, "Real-Time Lip Synchronization between Text-To-Speech (TTS) System and Robot Mouth", 19th IEEE International Symposium on Robot and Human Interactive Communication, 2010.