

RECONSTRUCT 3D HUMAN FACE USING TWO ORTHOGONAL IMAGES

H. S. Mazumdar¹, Priyank Y. Chauhan²

¹Professor and Head, Research and Development Center, Dharamsinh Desai University, Nadiad, Gujarat, India

²Research scholar, Department of Information Technology, Dharamsinh Desai University, Nadiad, Gujarat, India

Abstract

As the 3D human face reconstruction is becoming very popular in recent times, it attracts many researchers. Construction of 3D human face using only two orthogonal images and twelve landmark features are the main context of the proposed approach. For 3D object modeling, the Open Graphics Library (OpenGL) is used as the platform through which modeling, modification and rendering is performed on the morphable model. The proposed approach trails through semi-automatic identification of the facial landmark features, calculation of the 3D coordinates of human face, morphable model construction in OpenGL, reshaping of the morphable model and rendering of the morphable model. The facial landmark identification is semiautomatic method, as the module requires the manual interaction for marking the facial landmarks on the image. The reshaping of morphable model is required as the morphable model does not fit to the actual face in most of the cases. The morphable model is reshaped by calculating root mean square (RMS) error of face coordinates. The rendering process does not require a wide screen image because the approach performs rendering using input front face image and side face image as textures. Applications of this research help to overcome the disputes of fields like crime detection, 3D game characterization, the ornaments exhibition and in the area of medical technology like plastic surgery.

Keywords: Morphable model, Feature Identification, Face Reconstruction, Texture Mapping.

1. INTRODUCTION

Human head is difficult to model even using normal 3D modeling software [4] hence the reconstruction of human face is being studied in last few decades [5][9]. The humans can perceive the 3D shape of a 2D image by just looking at it, even if the object in the image is completely new to the eye. The human eye signals the brain about the object through a nerve signal and from this nerve signal, the brain creates the 3D shape of the 2D object [4]. The same process is used for 3D generation in real life. The knowledge of image processing can help to simplify the face recognition and allow prediction of appearance under novel viewing conditions. Human face identification is a challenging task due to uniqueness of human face. Landmark feature is a subset of face like eyes, nose, lips, ears, cheeks which has unique function to perform. Facial signature is the shape of face generated by combining all facial landmark together. Facial signature is generated by the landmark features through which human face can be identified.

The major applications of the concept are enlisted here. In real life, the photos taken in crowded places are not always front or side facing. In these photos, the people are in action and have different posture and illumination condition. The recent evidence states that the structural 3D information captured by face scans can improve face recognition particularly in those situations where pose variations and illumination changes are concerned [3]. So the face identification can be achieved in any situation effectively. The crime detection is the main application of the concept. 3D games can be made realistic by placing actual character's

face at cartoon character. 3D and animation movies also get such perks.

Techniques for attaining facial information for 3D reconstruction are broadly categorized into three techniques, namely, pure image-based techniques, hybrid image-based techniques and 3D scanning techniques. The pure image-based techniques perform the reconstruction using only 2D images without estimating the real 3D structure. In hybrid image-based techniques both approximations and the data gained from images are used in the reconstruction process. The 3D scanning techniques have the capability to capture the complete 3D structure since scanned images provide both, geometry and texture information of the face [4].

Section 2 will explain the background theory. Section 3 will explain the proposed approach. Section 4 will conclude our work.

2. BACKGROUND THEORY

The concept of 3D object modeling is carried out by the Stereoscopic effect. If the object is seen with only left eye opened, the object is a bit more of the left side same as right eye opened, the object is a bit more of the right side. The brain fuses the two images together allowing one to see in three dimensions. This is known as stereoscopic vision. The most known method to generate 3D is Shape-from-shading, which takes a single image as input and uses single 3D reference model of a different person's face [1]. 3D Face reconstruction is a method in which the facial features are identified and extracted using Radial Basis Functions (RBF).

The reconstructed model is in the Virtual Reality Model Language (VRML) format [10].

The morphable model [5] [6] [7] [9] is widely used in different methods as it provides transferring, reshaping and texturing processes to be possible with simplicity. The morphable model is taken as the base object and it can be manipulated, viewed and rendered in 3D space. The texture mapping requires a base image for rendering process that can be placed on the morphable model. One of the textures mapping algorithm like cubical ray projection algorithm [2] uses the linear equation to recover the texture and shape parameters irrespective of lighting conditions and the posture of the input face image.

3. PROPOSED APPROACH

The step by step methodology is described in details for more understanding. The preprocessing task requires the following steps.

1. Take 2D images of front and side face.
2. In front face image, mark the three landmarks namely left and right eyes and center of lips.
3. In side face image, mark the two landmarks namely left eye and center of lips.
4. In front face, crop the image by marking top, bottom, left and right boundaries with retaining 10% proportional area with face boundary.
5. In side face also, crop the image by marking top, bottom, left and right boundaries with retaining 10% proportional area with face boundary.
6. Normalize the front face by rotating it such that the median of two eyes and center of lips comes in vertical line.
7. From front face, take horizontal bars of eyes and center of lips. According to these horizontal bars perform translation and matching on eyes and center of lips of side face.

For aligning and normalizing the landmark features, it is necessary to take orthogonal images as input for preprocessing task. Process starts with marking the facial landmark on images followed by cropping and rotating.

For the front face, mark the three facial landmarks namely both eyes and center of lips. These are the facial landmarks which are visible in both the face images and used as aspect ratio to match them with the original face. For the side face, mark the two facial landmark namely eye and center of lips. Now the images have to be cropped such that the images can be used to identify the landmark features. To crop the front face image, first four boundaries has to be marked, top, bottom, left and right side boundaries. The cropped image must retain 10% proportional area with the boundary of human face in each side. Same way generate the cropped image of side face.

To normalize the front face, first find the median of the two eyes and then rotate the image until the median point and center of lips comes in the vertical line. To find the median of the two eyes take average of x coordinates of both eyes.

Now take one horizontal bar passing through eyes and another passing through center of lips. Horizontal bars are parallel to x-axis so the value of y coordinate remains same on the horizontal bar. Now take side face image and place the image such that the horizontal bars matches with the eye and center of lips. The horizontal bar passing through eyes of front face has to pass through the eye of side face and the same way for center of lips. Both the images are normalized now.

These are few assumptions which are followed by us in proposed approach to reconstruct 3D human face. The side face of human face is assumed to be symmetric on both the side. In other words, left side face is considered to be the same as the right side face of the human face. The accessories like ear-rings, glasses should not be worn while taking photos. The haircut or hairstyle shown in 3D face is ideal for all generated 3D face as each person has different hairstyle or haircut.

The proposed approach uses only twelve landmark features that are marked by human to generate unique facial signature. From that it produces 3D human face which is more likely suitable for visual analysis.

The following steps are performed to generate 3D human face.

1. The aligned and normalized outcome of preprocessing task is taken as input images.
2. Facial landmarks are identified manually by marking two eyes, two ears, two cheeks, center of lips, top of head, center of forehead, nose tip, chin in front face and eye, ear, cheek, center of lips, top of head, and center of forehead, nose tip, and chin and back of head in side face.
3. Calculate 3D coordinates for twelve facial landmarks using following equation.

$$(x, y, z) = \left(xf, \left(\frac{yf + ys}{2} \right), xs \right) \quad [10] \quad (1)$$

xf – X coordinate from front face image.

yf – Y coordinate from front face image.

xs – X coordinate from side face image.

ys – Y coordinate from side face image.

In case, the front and side view are registered using landmarks, yf and ys are very close to each other.

4. Create 3D morphable model in OpenGL.
5. Reshape the morphable model performing translation, rotation and scaling operations using only nine landmarks.
6. Manual editing for interior landmark features like eyes, nose, lips, ears, eyebrows.
7. The Texture Generation Algorithm is used to produce texture image for texture mapping.
8. The texture mapping utility of OpenGL is used for mapping texture images to morphable model.

In this novel method, the reconstruction process is divided into four steps namely identification of landmarks, construction of morphable model, reshaping of morphable model and texture generation and its mapping on morphable model. We are using Fig - 1 as a reference image for all input images. The semiautomatic identification is carried out using the reference image and color code mapping technique. In reference image, the facial landmarks are marked with color and the color code values are used to identify the landmarks. The face images are normalized in between 0 to 1 coordinate values.

Paragraph comes content here. Paragraph comes content here. Paragraph comes content here. Paragraph comes content here. Paragraph comes content here. Paragraph comes content here.

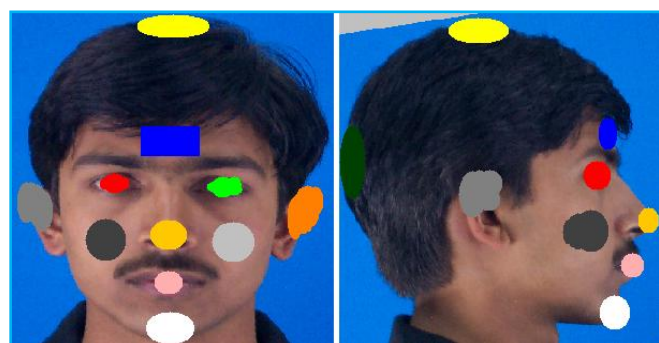


Fig -1: Reference Image.

The twelve landmarks are generated by combining eleven landmarks of front face and nine landmarks of side face those are marked manually. Fig - 2 shows all twelve landmarks which are top of head, center of forehead, left eye, right eye, left ear, left cheek, nose tip, right cheek, right ear, center of lips, chin and back of head. The mismatched landmarks are back of head, eye, ear and cheek in both the images.



Fig -2: Front face and Side face landmark Features.

Feature identification is performed manually on the human face by above approach but it is also possible to automate this process by projecting chess pattern on human face at 45° angle and locating the corners on the front and side images.

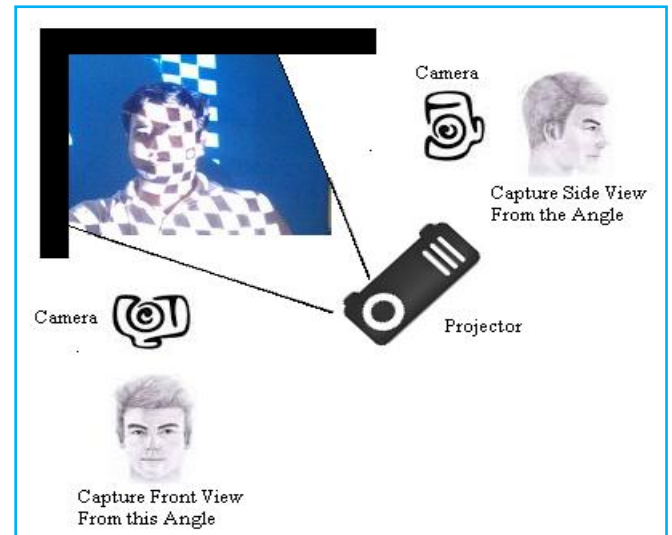


Fig -3: Camera Projector Setup.

Fig - 3 shows the camera projector set up. In chess pattern images the black squares are replaced with white squares in alternate image. The images are projected on the face as per shown in the Fig - 3.

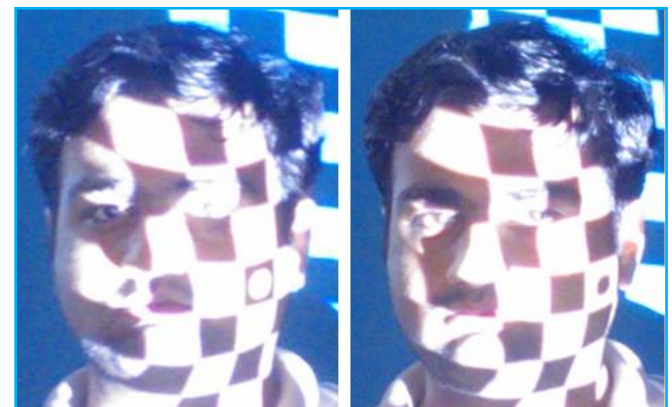


Fig -4: Front view image of Chess Patterned Human Face.

Fig - 4 and 5 shows the front and side view image respectively with dot being located on the cheek and other feature points are searched in reference to it by locating square corners.



Fig -5: Side view image of Chess Patterned Human Face.

To automate landmark identification, following steps has to be performed.

1. Locate the dot in the front and side view.
2. Find the respective corners of the black squares in front and side view.
3. Match them with morphable model.

The paper trails through manual landmark identification method as automatic method is yet to implement.

The object has [width, height] property in 2D. And for the 3D space, the object has an extra property called depth and represented by [width, height, depth] of the object. Width of the object in 3D space is measured from front face only; depth is measured from side face only and the height of the object from both the faces. When 3D object is plotted into 2D space, the front face image produces [width, height] properties and side face produces [depth, height] properties. As per these two concepts, the 3D coordinates' x value is x_f , z value is x_s and y value is $\left(\frac{y_f+y_s}{2}\right)$. These 3D coordinates adjust the morphable model's landmarks into 3D space. The morphable model has appearance like human face as shown in Fig - 6.

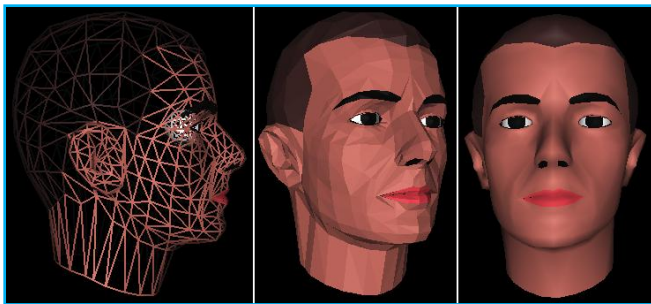


Fig -6: Morphable model constructed in OpenGL.

The morphable model and the input images have very less possibility to be similar that's why it is necessary to reshape the morphable model. The Center of Gravity (CG) and Root Mean Square (RMS) are used for reshaping of morphable model.

Center of Gravity: it is the center point of all the landmark points. Operations like translation, scaling, rotation are performed by taking CG as reference.

Root Mean Square: it calculates distance between CG and each landmark points.

The steps involved in reshaping of morphable model are described as below.

1. Calculate CG of front face image and morphable model only for front face landmark points excluding both cheeks landmarks.
2. Calculate RMS distance of remaining nine landmark points with respect to CG for morphable model and front face image.
3. Perform translation, rotation, translation, scaling, translation, rotation and scaling operation

sequentially on the landmark points with respect to CG.

4. Repeat step 3 until RMS distance reduces to least and constant.
5. Similarly repeat step 1 to step 4 for side face image excluding cheek landmark.

The reduced RMS distance shows the optimum matching of morphable model to the actual face. The cheeks are excluded for reshaping algorithm as first objective is to reshape morphable model to match it with boundary of the face. The use of cheek landmarks is for automatic texture mapping algorithm. The actual morphable model is shown in Fig - 6. After performing reshape, the morphable model is modified and it looks like as shown in Fig - 7 and 8. Fig - 7 is in solid fill and Fig - 8 is textured mode. The morphable model is shown with the front and side images for comparing and mapping. The view is taken from 45° and -45° angle. These angles are best suited to match and verify 3D and 2D. The images when placed to the center, front and side face matches with each other. But for clear view to ears from front, the front face image is set at some distance.

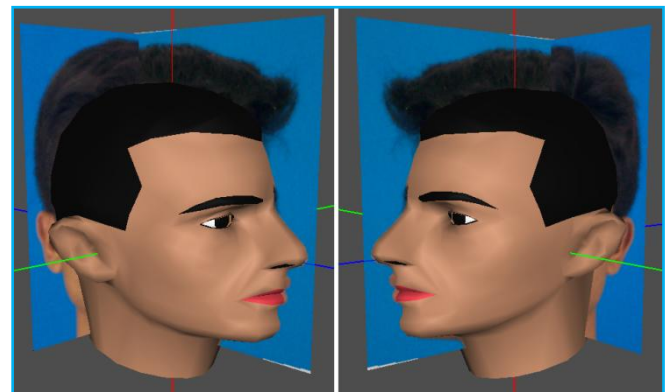


Fig -7: 45° and -45° angle view of 3D solid morphable model.

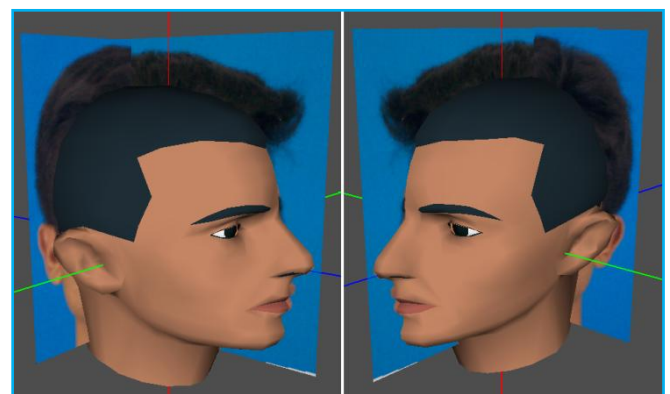


Fig -8: 45° and -45° angle view of 3D textured morphable model.

Reshape algorithm matches front and side face of the morphable model with the input face boundary. Fig - 9 and 10 shows the exact matching of morphable model and actual face boundary.

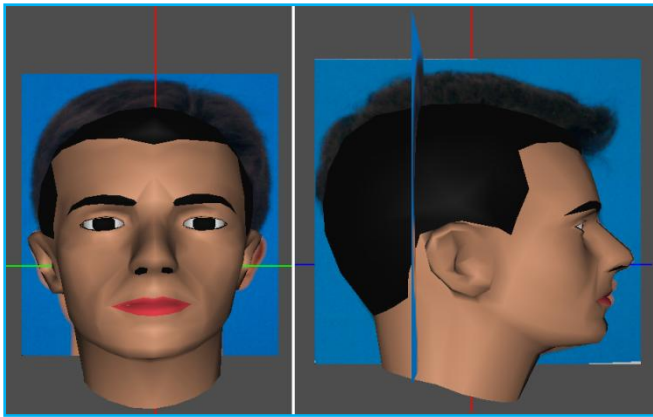


Fig -9: Front and side view matching.

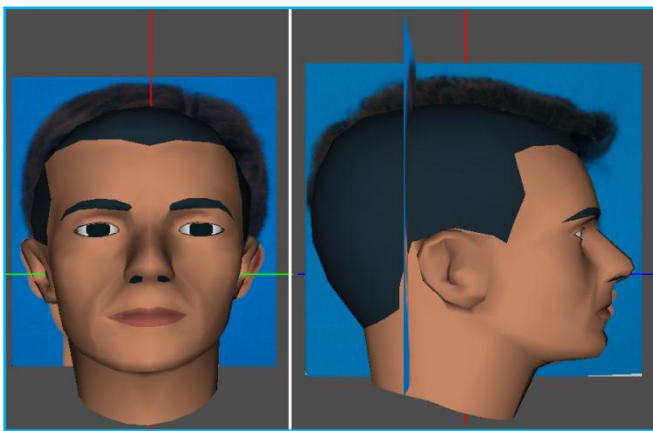


Fig -10: Front and side view matching.

Front Face and Side face boundary matching is achieved with very much accuracy. Fig - 11 shows the accuracy of matching. Still some manual matching of interior landmark has to be performed for higher accuracy.

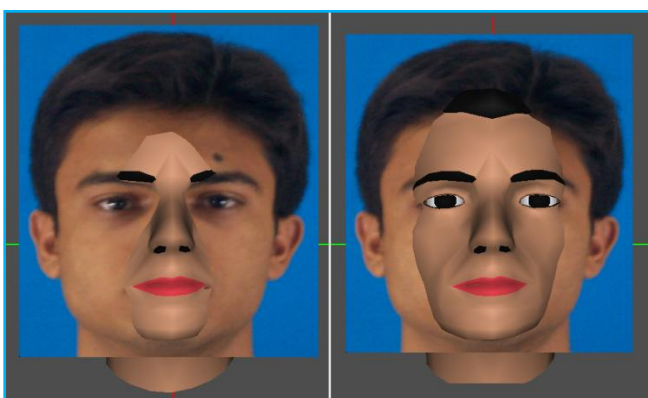


Fig -11: Front face and landmark matching.

After boundary matching, manual matching of interior landmarks is performed. The interior landmarks are like eyes, lips, ears, nose, eyebrows etc. Fig - 12 shows the matching of interior landmarks like nose, ears. Other interior landmarks are also matched manually. The automatic way to match

these landmarks is hard to define, yet not impossible. This can be achieved, but not in such short time.

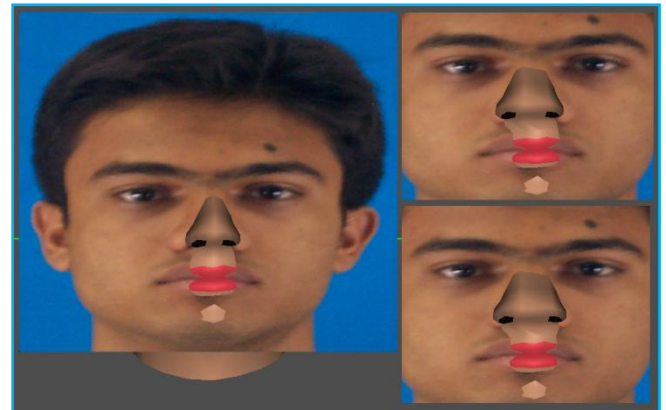


Fig -12: Manual nose matching.

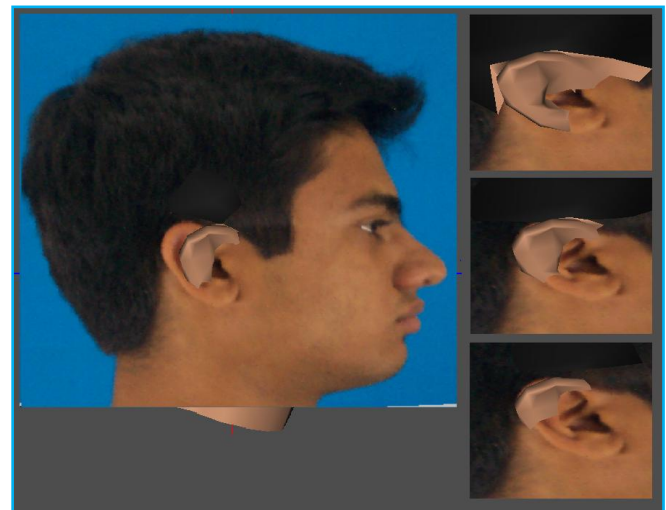


Fig -13: Manual ear matching.

OpenGL provides inbuilt support for texture generation and texture mapping provided with normalized texture coordinates. Using OpenGL, it is easy to model and to render any OpenGL object [12].

Texture can refer to the actual image and it adds details to the computer generated graphics. Texture mapping is useful to get more realistic face [8].

Texture generation and mapping is carried out in following way.

1. Calculate triangle area for each and every triangle of morphable model by using (x, y) coordinates and (y, z) coordinates.
2. The higher one decides from which image the texture portion has to be taken for texture mapping.
3. The morphable model vertex specifying particular landmark is passed with the stored texture coordinate.
4. Generate texture coordinates for nearby vertexes with respect to landmark vertexes.

- Calculate texture coordinates from front and side images.

Texture coordinates are passed to OpenGL to generate texture. The stored value of landmark features is normalized in-between 0 to 1 and same is used for texture coordinates. Fig - 8, 10 and 15 are textured images. Fig - 14 shows morphable model in solid fill, wireframe and textured mode.

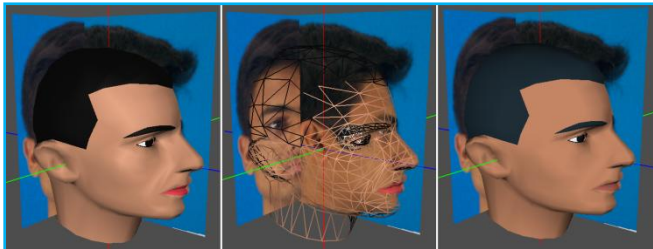


Fig -14: Solid, Wireframe and Textured morphable model.

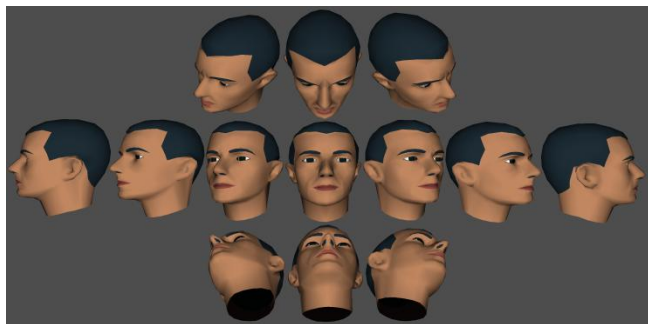


Fig -15: Textured morphable model from all the angles.

This is fully automatic approach as no user interaction is required. The approach is still under development. The texture taken for texturing is one point texture and is taken only from front face image. In this, the skin texture is taken from the cheek, lip texture is taken from lips and hair, eye center and eyebrow texture is taken from top of head.

One point texture: The generated texture image is taken from only one pixel. The image from the pixel is stretched to make it as the shape of the triangle.

4. CONCLUSIONS

It is possible to reconstruct the 3D human face using existing morphable model, as the name suggests reconstruction. The proposed approach overcomes various difficulties like selection and manipulation of too many landmark features, number of input images, generation of widescreen texture image, requirement of gadgets and identification of landmarks manually. Generation using only two orthogonal images; front face and side face and twelve landmark features draw researchers' attention. Manual modification of interior landmark improves accuracy but takes increases time. Calculation of triangle area takes very less time as the number of triangles used in morphable are 1509. Only texture generation and mapping uses cheeks to improve accuracy of rendering. Accuracy can be achieved

most by marking accurate landmark. The final result is comparable with the results of available tools. In future, it would be interesting to see whether the approach helps for human body reconstruction also.

ACKNOWLEDGEMENTS

I convey my sincere thanks to the Head of the Department of the Information Technology Prof. R. S. Chhajed for kindly providing resources whichever needed for my dissertation and also allowing me to research and complete my dissertation at R&D center. I am also thankful to Mr. Himanshu Purohit and Mr. Marichi Patel for their help whenever I needed it.

I convey my sincere thanks to DDU, for providing me the database of images of DDU staff for my dissertation research.

I am grateful to Dr. Steve Dipaola and his team who developed the morphable model in OpenGL using C (COGL). Their work inspired and motivated me to carry out the research work. Without their work, it would have been very hard me to complete the research work.

REFERENCES

- [1] Ira Kemelmacher-Shlizerman, Ronen Basri, "3D Face Reconstruction from a Single Image using a Single Reference Face Shape", IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE.
- [2] M. RamaSubramanian, M.A. Dorai Rangaswamy, "Reconstruction of sensible 3D Face Counterpart From 2D Images - A hybrid approach", International Journal of Computational Engineering Research (IJ CER) ISSN: 2250-3005, National Conference on Architecture, Software system and Green computing.
- [3] Stefano Berretti, Alberto Del Bimbo, Pietro Pala, "3D Face reconstruction from two orthogonal Images for Face recognition Applications", International Journal of Digital Library Systems, 1(3), 42-58, July-September 2010.
- [4] W.N. Widanagamaachchi, A.T. Dharmaratne, "3D Face Reconstruction from 2D Images A Survey", Digital Image Computing: Techniques and Applications, University of Colombo School of Computing, 35, Reid Avenue, Colombo 7, Sri Lanka.
- [5] Ateş Akaydın, Onur Küçüktunç, "3D Face Reconstruction from 2D Images for Effective Face Recognition", Bilekent University, Department of Computer Engineering, Bilkent, Ankara, Türkiye.
- [6] Volker Blanz, Thomas Vetter, "A Morphable Model For the Synthesis Of 3D Faces", Max-Planck-Institute for biologist Kybernetik, Tubingen, Germany.
- [7] Sami Romdhani, Volker Blanz, and Thomas Vetter, "Face Identification by Fitting a 3D Morphable Model Using Linear Shape and Texture Error Functions", "Springer-Verlag Berlin Heidelberg

- 2002”, University of Freiburg, Institut for Informatik, Freiburg, Germany.
- [8] WonSook. Lee, Jin Gu, and Nadia Magnenat, “Generating Animatable 3D Virtual Humans from Photographs”, “EUROGRAPHICS 2000”, MIRALab, CUI, University of Geneva, Switzerland.
- [9] Changhu Wang, Shuicheng Yan, Hongjiang Zhang, Weiyang Ma, “Realistic 3D Face Modeling by Fusing Multiple 2D Images”, “Proceedings of the 11th International Multimedia Modeling Conference 2005 IEEE”, Microsoft Research Asia, Beijing.
- [10] Nikhil Rasiwasia, Dr. K S Venkatesh, “3-D Face reconstruction from two orthogonal pictures with application to facial makeover”, Dept of Electrical Engineering, IIT Kanpur.
- [11] [Façade – Stanford Facial animation system:] <http://www.dipaola.org/stanford/facade/presentation.html>.
- [12] Dave Shreiner, Graham Sellers, John Kessenich, Bill Licea-Kane, “OpenGL Programming Guide: The Official Guide to Learning OpenGL, Version 4.3. Eighth Edition”.