KLE Society's
KLE Technological University

A Minor Project Report

On

MOBILE APP FOR 3D FACE GENERATOR

submitted in partial fulfillment of the requirement for the degree of

Bachelor of Engineering in

Computer Science and Engineering

Submitted by:

NITISHA SINHA	**01FE16BCS130**
JYOTHI S HOSAMANI	**01FE16BCS083**
SHRIYA HIREHOLI	**01FE16BCS079**
ISHIKA KUMAR	**01FE16BCS081**

Under the guidance of
Dr. Shrinivas D. Desai

.

SCHOOL OF COMPUTER SCIENCE & ENGINEERING

HUBLI – 580 031 (India).

Academic year 2018-19

KLE Society's
KLE Technological University

2018 - 2019

SCHOOL OF COMPUTER SCIENCE & ENGINEERING

CERTIFICATE

This is to certify that Minor Project entitled **Mobile App for 3D Face Generator** is a bonafied work carried out by the student team Ms. Nitisha Sinha- 01FE16BCS130, Ms. Shriya Hireholi – 01FE16BCS079, Ms. Jyothi S Hosamani - 01FE16BCS083, Ms. Ishika Kumar- 01FE16BCS081, in partial fulfillment of completion of Sixth semester B. E. in Computer science and Engineering during the year 2018 – 2019. The project report has been approved as it satisfies the academic requirement with respect to the project work prescribed for the above said program.

Guide — **Head of SoCSE**

Dr. Shrinivas D Desai — **Dr. Meena S. M**

External Viva:

Name of the Examiners — **Signature with date**

1.

2.

1. INTRODUCTION

In this chapter, we are going to present the what exactly is 3D face reconstruction, several techniques to implement the same and the existing market trend of 3D face reconstructing model as a product. We will also be seeing what motivated us to take up this project and get the overview of the same which led us to identify the domain, come with a concrete problem statement and the objectives.

In computer vision and computer graphics, **3D reconstruction** is the process of capturing the shape and appearance of real objects. This process can be accomplished either by active or passive methods. If the model is allowed to change its shape in time, this is referred to as non-rigid or spatio-temporal reconstruction. The research of 3D reconstruction has always been a difficult goal. Using 3D reconstruction one can determine any object's 3D profile, as well as knowing the 3D coordinate of any point on the profile. The 3D reconstruction of objects is a generally scientific problem and core technology of a wide variety of fields, such as Computer Aided Geometric Design (CAGD), computer graphics, computer animation, computer vision, medical imaging, computational science, virtual reality, digital media, etc. For instance, the lesion information of the patients can be presented in 3D on the computer, which offers a new and accurate approach in diagnosis and thus has vital clinical value. Digital elevation models can be reconstructed using methods such as airborne laser altimetry or synthetic aperture radar.

3D reconstruction from multiple images is the creation of three-dimensional models from a set of images. It is the reverse process of obtaining 2D images from 3D scenes.

The essence of an image is a projection from a 3D scene onto a 2D plane, during which process the depth is lost. The 3D point corresponding to a specific image point is constrained to be on the line of sight. From a single image, it is impossible to determine which point on this line corresponds to the image point. If two images are available, then the position of a 3D point can be found as the intersection of the two projection rays. This process is referred to as triangulation. The key for this process is the relations between multiple views which convey the information that corresponding sets of points must contain some structure and that this structure is related to the poses and the calibration of the camera.

In recent decades, there is an important demand for 3D content for computer graphics, virtual reality and communication, triggering a change in emphasis for the requirements. Many existing systems for constructing 3D models are built around specialized hardware (e.g. stereo rigs) resulting in a high cost, which cannot satisfy the requirement of its new applications. This gap stimulates the use of digital imaging facilities (like a camera). Moore's law also tells us that more work can be done in software. An early method was proposed by Tomasi and Kanade.They used an affine factorization approach to extract 3D from images sequences. However, the assumption of orthographic projection is a significant limitation of this system.

There exist **several techniques** to construct a 3D representation of a subject. Computed Tomography (CT)-scans and Magnetic Resonance, for example, allow for easy 3D surface extraction in more demanding medical scenarios, but due to either significant radiation exposure or high costs respectively, they are not acceptable for aesthetic procedures. In addition, no texture information is present on these types of data. Other expensive and cumbersome hardware dependent solutions such as 3D scanners based on stereo cameras or laser are also available, but only few centres have access to these costly devices that are, in addition, complex to use. The data provided by 3D scanners are typically not ready to use and require intensive post-processing of specialized personnel. Moreover, hardware dependent solutions typically hinder support and maintenance, as well as hinder the implementation of web-based applications that can reach a wider range of users. Alternatively, hardware independent 3D face creation has been explored since many decades in other areas following the pioneer work of Parke using a parametric model of the human face. In such approaches, 3D faces are created without using special hardware such as multiple calibrated cameras or projecting devices. In this context, many different techniques have been proposed including shape from shading structure from motion and statistical facial models. Each of these techniques have advantages and disadvantages. Shape from shading can provide good and smooth facial shape for very controlled environments and with well-known light sources and models. However, it is very sensitive to light variation, which cannot be controlled from one clinical scenario to the other. Structure from motion is less sensitive to the light conditions, but is sensitive to the variations on camera point of view hindering implementation into clinics. Statistical facial models have shown to be very robust and to allow fast creation of 3D faces by adapting the variation stored in the model to an image of a specific face. Previous work has already showed the potential of techniques based on statistical facial model in clinics. In for example, the authors show that the accuracy achieved with a similar method (face reconstruction based on three images) is compatible with the accuracy expected by clinicians for a communication tool in aesthetic procedures. However, limited accuracy has been reported in certain regions (e.g. cheeks area), since the method used is based on images seen from only three points of view not containing information about those regions. Although a very effective tool could be created, the methods can still be further improved to better represent the patients. In summary, current hardware independent solutions are, in certain extent, limited in either speed, robustness or accuracy, and therefore leave space for investigation of new techniques.

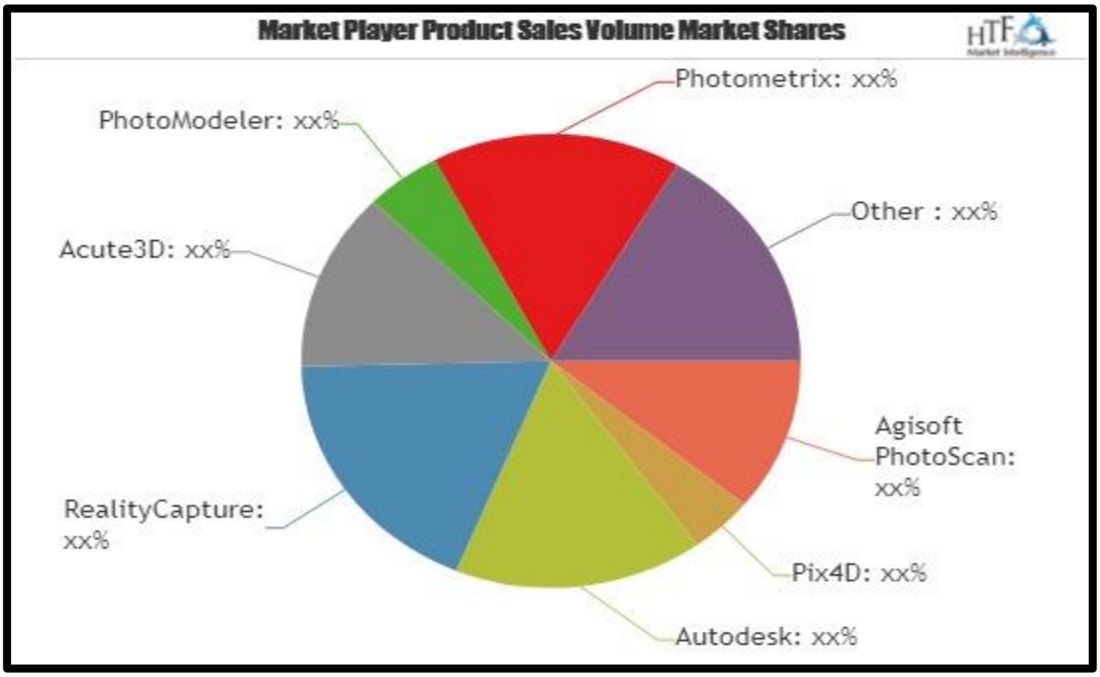

Fig.1(a) Depicts a pie chart of the product sales volume market shares in 3D face reconstruction/3D modelling.

1.1 MOTIVATION

- Customized 3D view of real time images will be more appealing to children of age group 5-13 years.
- Academic materials in this 3D format will have more impact in teaching as well as learning process.
- There are many apps/products available in the online market which are having limitations like single “front view” less of “comic” in nature.
- Hence, this motivated us to develop interactive 3D view and more realistic representation of given 2D image.

1.2 OVERVIEW OF THE PROJECT

Capitalizing on the advances of 3D face reconstruction from computer graphics and computer vision, this work presents a system to enable precise and robust 3D facial reconstruction from video. The reconstructed face can be used for emulation of aesthetic procedures in 3D. Depending on the setting and the assumptions made, there are many variations of it as well as a multitude of approaches to solve it. The proposed 3D face reconstruction algorithm uses a statistical shape model as well as information from facial landmarks and silhouette to iteratively model the face of a subject moving the face from front to left. The presented 3D active shape model approach enables the spatio-temporal tracking of facial landmarks and pose estimation based on a few initial facial landmarks defined in the first frame. Silhouette information extracted from key frames allows for better face reconstruction.

Module One:

Title: Mobile App for Face Cropping and Face expression generator.

Requirements: Whenever user provide an image as input the face present in that image should be cropped exactly and then applying different image wrapping or any other technique exact expressions such as smile, sad, crying, angry laughing and pale should be generated.

Module Two: Mobile App for Face 3D Generator.

Requirements: Input the single image of the human face. Generate a 3D model of their face.

Module Three:

Title: Mobile App to convert input Face into a Cartoon Bobble head.

Requirements: input the whole 3D face and convert the face into a cartoon bobble head i.e. the given face, a cartoon face should be generated by applying different shades to input image.

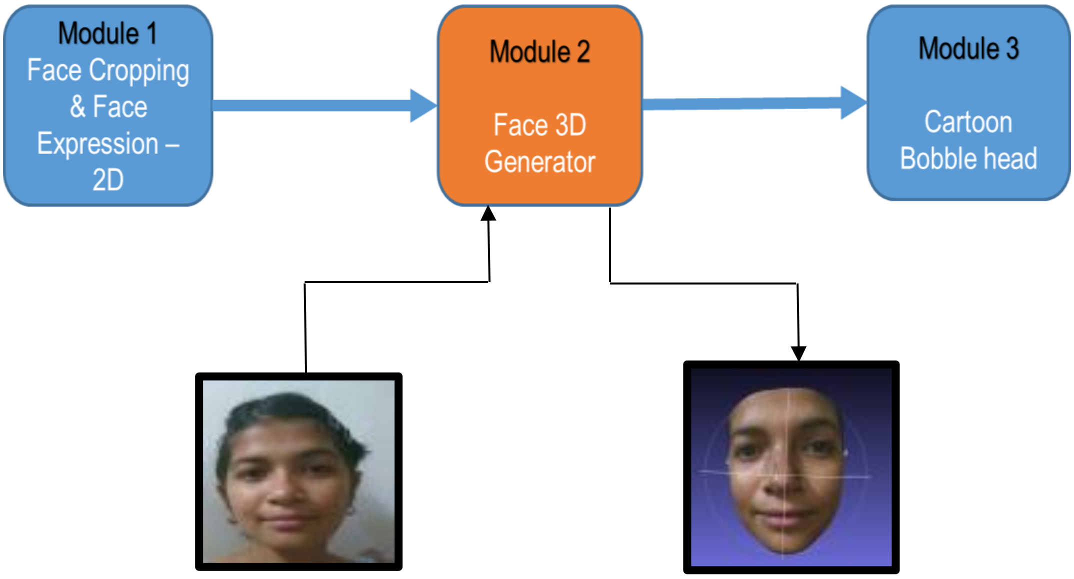

Fig.1.2(a) Depicts the pipeline of the entire system and highlights the input and expected output of the Module 2.

1.3 DOMAIN IDENTIFICAION

Computer Vision

Computer vision is concerned with the automatic extraction, analysis and understanding of useful information from a single image or a sequence of images. It involves the development of a theoretical and algorithmic basis to achieve automatic visual understanding.

Construction of 3D face using single image as an input

Constructing 3D face of the user from the single image as an input.

1.4 PROBLEM STATEMENT

To develop mobile/system friendly app, which will generate 3D face with 180° rotation by using video as input.

1.5 OBJECTIVES

- To read the single image of the human face.
- To represent 3D view of input face image.
- To develop the solution which should be integrated with other modules 1 and 3.

In the next chapter, we are going to know about literature survey done.

2. LITERATURE SURVEY

In this chapter, we will be knowing what is literature survey and the significance it poses before undertaking any project. It also presents with the literatures of various authors surveyed based on the techniques used, their advantages, programming platforms used, issues faced, etc. that helped during the implementation of the given project. We will also see the challenges that were addressed during the same.

Many algorithms have undergone development for the conversion of 2D images to 3D models. While many of these algorithms handle inanimate objects without much variation, others focus primarily on reconstructing the animate objects such as the human face, which offers a lot of variety, in 3 dimensions. Different researchers have offered a wide variety of solutions to this problem.

In [1], the researchers have condemned the inefficient and complex pipelines for model building and fitting. Instead, a Convolutional neural network is fed with 2D images and 3D facial scans. This CNN architecture performs direct regression of a volumetric representation of the 3D facial geometry from a single 2D image. In [2], a new technique called the 3DMM or 3D Morph able Model is deployed for estimating the full 3D facial structure from a single image. Two hundred laser-scanned heads were used as input data and each input was broken down into segments, varying the parameters of which produced a variety of new faces from the same base input. In [3], the researchers propose a novel method for 3D shape reconstruction of faces from a single image by using only a single reference model. This method exploits the global symmetry of faces by combining the shading information. In [4], the input is a video of a person's face and each frame is then reconstructed into a high detail 3D shape. This is made possible with the help of a new dense 3D flow estimation method coupled with shape from shading. In [5], the Caffe deep learning framework is used to train the DNN models on both real 2D images as well as synthetic 2D images. The final product is the 3D model capturing the facial expression as well.

Expression generation for an image is also a considerably challenging task. The researches in this area depends largely on the researcher's perspective. Some adopt an anatomical approach and study the underlying relations between bones and muscles for generating expressions. In [6], the researchers present a framework, which can generate expressions based on the active contour technique coupled with mesh warping algorithm (for morphing). In [7], the researchers use local warping methods for facial expression warping. The deformations are handled using with the help of two transformations namely, piecewise polynomial transformation and radial basis function transformation.

Table 2(a) Literature survey for 3D face reconstruction

Sl. no	**Author name**	**Proposed technique/ tool**	**Dataset**	**Program Platform**	**Issues**	**Remarks**
1.	Christian Baumberger, et.al, 2016	3D Face Reconstruction from Video Using 3D Morphable Model and Silhouette	**1**. in-model registered (IMR) **2**. out-model registered (OMR) **3**. out-model nonregistered (OMNR)	---		
2.	Christopher B. Choy, et.al, 2016	3D-R2N2: A Unified Approach for Single and Multi-view 3D Object Reconstruction	**1.** PASCAL 3D **2.** ShapeNet **3.**Online Products **4.**MVS CAD Models	---	**1.** It cannot reconstruct many details when given more than 30 different views of the model. **2.** Worst performance in reconstructing objects with high texture levels.	**1**. Does not require any image annotations (i.e., no segmentations, key points). **2.** It can selectively update hidden representations by controlling input gates.
3.	Aaron S. Jackson, et.al, 2017	Large Pose 3D Face Reconstruction from a Single Image via Direct Volumetric CNN Regression	**1.**AFLW2000 - 3D **2**.BU-4DFE **3**.Florence	---	The performance of our method decreases as the pose increases.	---
4.	Wael AbdAlmageed, et.al, 2016	Face Recognition Using Deep Multi-Pose Representations	**1.** CASIA-WebFace **2.** IARPA Janus Benchmark-A (IJB-A) **3.** Labeled Faces in the Wild (LFW)	---	---	This multi-pose representation-based recognition pipeline outperforms other state-of-the-art methods as the algorithm is data-agnostic.

5.	Tomoya HARA, et.al, 2012	Fast-Accurate 3D Face Model Generation Using a Single Video Camera	---	---	---	1. Only using a single consumer video camera. 2. Computational-inexpensive and relatively-robust. 3. Fully-automatic 4. Can maintain a human-like face structure. 5. Can obtain more rich information from an input image sequence.
6.	Joseph Roth, et.al, 2016	3DMM (3D Morphable Model) to create a personalized template which adaptively influences reconstruction in a coarse-to-fine scheme.	M001 of the BU-4DFE database	C++ and MATLAB	Needs large number of images as an input for accurate results	---
7.	Jaeik Jo, et.al, 2015	Structure from Motion (SfM) and the 3D Morphable Model (3DMM)	3D facial scans from 150 subjects comprising 86 males and 64 females were obtained by a 3D laser scanner	---	Incurs a high computational cost due to fusion of SfM and S3DMM,	Person-specific 3D face can be reconstructed from only a single-view facial image, Robust to pose variations
8.	Pengfei Dou, et.al, 2017	fusion-CNN	real 2D images and synthetic 2D images to train the deep neural network	---	---	It is end-to-end and takes only a single RGB image as input
9.	Ruiqi Zhao, et.al, 2016	Feed-forward neural network (with 6 layers)	BP4D-S database (test dataset)	Python	Can be efficiently performed with	---

					large datasets	
10.	Luo Jiang, et.al, 2018	3D face recognition with Geometry details Fom single Image	FaceWarehouse and BFM2009	C++	Due to limited degrees of freedom of the low dimensional model, these methods often fail to reproduce fine geometric details (such as wrinkles) that are specific to the target face.	Combines the benefits of low-dimensional face models and shape-from-shading, enabling more accurate and robust reconstruction.
11.	Nikolai Chinaev, et.al, 2018	3D Face Reconstruction with Efficient CNN Regression	Face Warehouse and BU4DFE	Python	The model can’ generate occlusions and extreme lighting conditions.	---
12.	Amin Jurabloo, et.al, 2016	A Survey of different 3D Face Reconstruction Methods	---	---	In Multiview stereo, it can estimate the 3D shape of an arbitrary object but it needs multiple images with known camera views.	In SFM The disadvantage of this method is that the estimated 3D shape bases are person specific.

2.1 CHALLENGES

- The network occasionally makes mistakes when faced with facial attributes which were not present in the synthetic data.
- Due to limited degrees of freedom of the low dimensional model, these methods often fail to reproduce fine geometric details.

In the next chapter, we are going to see the system proposed.

3. PROPOSED SYSTEM

The proposed system gives an idea about the process performance. It also indicates the completeness and quality of the process. **The following section** describes the system model, the target users, applications of the project and the scope of present work.

3.1 SYSTEM MODEL

In order to understand how to use models to construct system and conceptualize it, we draw system model. A system model is used to describe and represent multiple views such as requirement analysis, design, implementation, behavior, input data and output data.

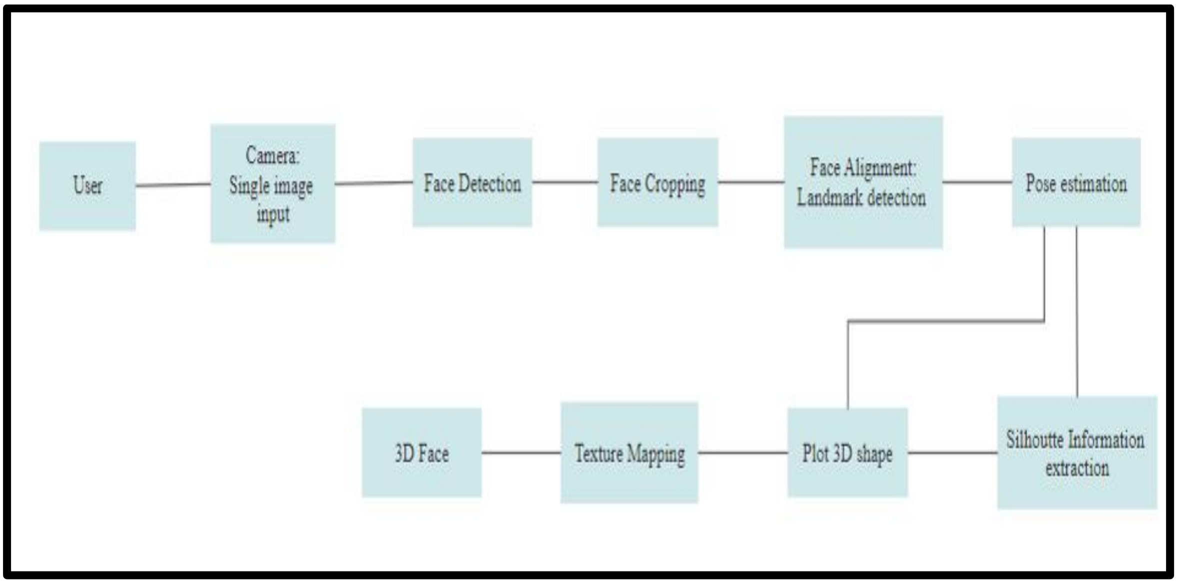

Fig.3.1(a) Depicts the system models involved in 3D face reconstruction system

3.2 DESCRIPTION OF TARGET USERS

The company’s content has always been child specific and thus, the beneficiaries would be the preexisting and newly added children and the parents, specific to whom the content will be developed.

3.3 SCOPE

The current content is mostly 2D graphics with no proper orientation of the characters. Thus, the undergoing project is undertaken with the idea of revolutionizing the company's content to enhance its appeal which in turn shall help them deliver better, more interactive and more specific content to its customers.

In the next chapter, we are going to know about the requirement analysis.

4. REQUIREMENT ANALYSIS

In this chapter, we are going to know about the phases involved in requirement analysis and how the requirements' quality can be improved. We will also learn about the identified functional and non-functional requirements; and hardware and software specifications of the given project.

In system engineering and software engineering, **requirements analysis** encompasses those tasks that go into determining the needs or conditions to meet for a new or altered product or project, taking account of the possibly conflicting requirements of the various stakeholders, analyzing, documenting, validating and managing software or system requirements. Requirements analysis is critical to the success or failure of a systems or software project. The requirements should be documented, actionable, measurable, testable, traceable, related to identified business needs or opportunities, and defined to a level of detail sufficient for system design.

Conceptually, requirements analysis includes three types of activities:

- Eliciting requirements: (e.g. the project charter or definition), business process documentation, and stakeholder interviews. This is sometimes also called requirements gathering or requirements discovery.
- Analyzing requirements: determining whether the stated requirements are clear, complete, consistent and unambiguous, and resolving any apparent conflicts.
- Recording requirements: Requirements may be documented in various forms, usually including a summary list and may include natural-language documents, use cases, user stories, process specifications and a variety of models including data models.

Requirements analysis can be a long and tiring process during which many delicate psychological skills are involved. Large systems may confront analysts with hundreds or thousands of system requirements. New systems change the environment and relationships between people, so it is important to identify all the stakeholders, take into account all their needs and ensure they understand the implications of the new systems. Analysts can employ several techniques to elicit the requirements from the customer. These may include the development of scenarios (represented as user stories in agile methods), the identification of use cases, the use of workplace observation or ethnography, holding interviews, or focus groups (more aptly named in this context as requirements workshops, or requirements review sessions) and creating requirements lists. Prototyping may be used to develop an example system that can be demonstrated to stakeholders. Where necessary, the analyst will employ a combination of these methods to establish the exact requirements of the stakeholders, so that a system that meets the business needs is produced. Requirements quality can be improved through these and other methods

- Visualization. Using tools that promote better understanding of the desired end-product such as visualization and simulation.
- Consistent use of templates. Producing a consistent set of models and templates to document the requirements.
- Documenting dependencies. Documenting dependencies and interrelationships among requirements, as well as any assumptions and congregations.

4.1 FUNCTIONAL REQUIREMENTS

Functional requirements explain what has to be done by identifying the necessary task, action or activity that must be accomplished. Functional requirements analysis will be used as the top-level functions for functional analysis.

The functional requirements as identified in the project given are as follows:

- Construct 3D face using video as the input.
- The generated 3D face should be rotated in range of 180° from left to right.
- Rendering should be smooth.

4.2 NON-FUNCTIONAL REQUIREMENTS

In systems engineering and requirements engineering, a **non-functional requirement** (NFR) is a requirement that specifies criteria that can be used to judge the operation of a system, rather than specific behaviors.

Non-functional requirements are often called "quality attributes" of a system. Other terms for non-functional requirements are "qualities", "quality goals", "quality of service requirements", "constraints", "non-behavioral requirements", or "technical requirements". Informally these are sometimes called the "ilities", from attributes like stability and portability. Qualities—that is non-functional requirements—can be divided into two main categories:

- Execution qualities, such as safety, security and usability, which are observable during operation (at run time).
- Evolution qualities, such as testability, maintainability, extensibility and scalability, which are embodied in the static structure of the system.

The non-functional requirements as identified in the project given are as follows:

- Appearance should be more realistic.
- It should be at least 90% accurate.
- 3D face should be generated within a minute.

4.3 SOFTWARE AND HARDWARE REQUIREMENT SPECIFICATIONS

Software requirements:

- **Operating system:** Windows, Linux,
- **Programming language:** Python, Java
- **Tools:** OpenCV, Dlib, TensorFlow

In the next chapter, we will be seeing about system design.

5. SYSTEM DESIGN

In this chapter, we present about system design and how it proves to be beneficial during implementation phase. This section also describes about the high-level diagram, the architecture diagram, the detailed design of each module involved in the present work.

Systems design implies a systematic approach to the design of a system. It may take a bottom-up or top-down approach, but either way the process is systematic wherein it takes into account all related variables of the system that needs to be created—from the architecture, to the required hardware and software, right down to the data and how it travels and transforms throughout its travel through the system. Systems design then overlaps with systems analysis, systems engineering and systems architecture.

5.1 HIGH-LEVEL DIAGRAM

High-level design (HLD) explains the architecture that would be used for developing a software product. The architecture diagram provides an overview of an entire system, identifying the main components that would be developed for the product and their interfaces. The HLD uses possibly nontechnical to mildly technical terms that should be understandable to the administrators of the system. In contrast, low-level design further exposes the logical detailed design of each of these elements for programmers.

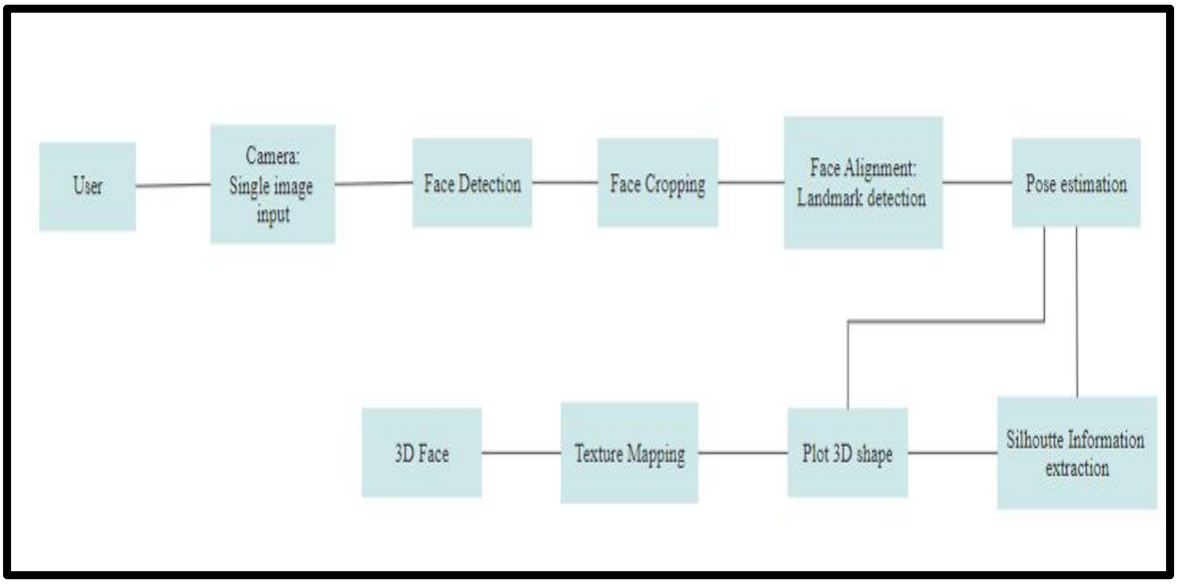

Fig.5.1(a) High level diagram which depicts the overall functioning of the 3D face reconstruction system.

5.2 ARCHITECTURE DIAGRAM

For system developers, they need system architecture diagrams to understand, clarify, and communicate ideas about the system structure and the user requirements that the system must support. It's a basic framework can be used at the system planning phase helping partners understand the architecture, discuss changes, and communicate intentions clearly. Using architecture diagram, you can also describe patterns that are used throughout the design.

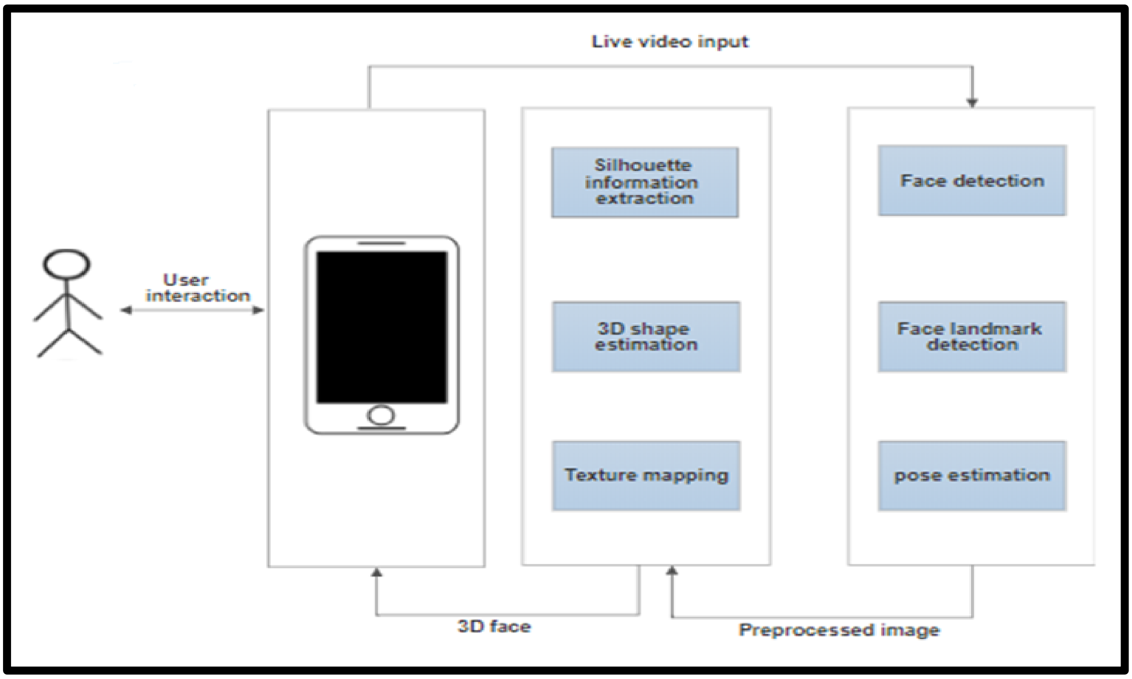

Fig.5.2(a) Depicts the architecture diagram of the 3D face reconstruction system.

5.3 DETAILED DESIGN

Detailed design of the system is the last design activity before implementation begins. The hardest design problems must be addressed by the detailed design or the design is not complete. The detailed design is still an abstraction as compared to source code, but should be detailed enough to ensure that translation to source is a precise mapping instead of a rough interpretation. The detailed design should represent the system design in a variety of views where each view uses a different modeling technique. By using a variety of views, different parts of the system can be made clearer by different views. Some views are better at elaborating a system states whereas other views are better at showing how data flows within the system. Other views are better at showing how different system entities relate to each through class taxonomies for systems that are designed using an object-oriented approach.

5.3.1 LANDMARK DETECTION MODULE

The landmark detection and pose estimation receive frames from a video and a set of initial landmarks defining facial features as input, and returns the location of a set of landmarks and the shape poses relative to the camera in each frame.

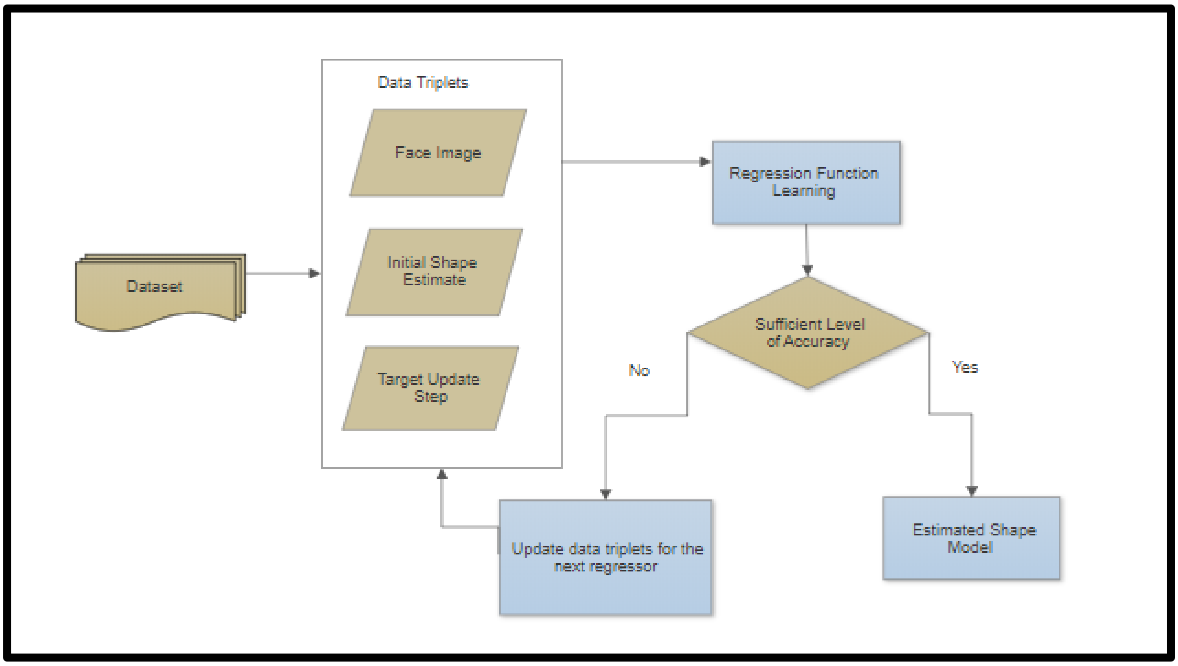

Fig.5.3.1(a) Depicts the process of landmark detection.

5.3.2 POSE ESTIMATION MODULE

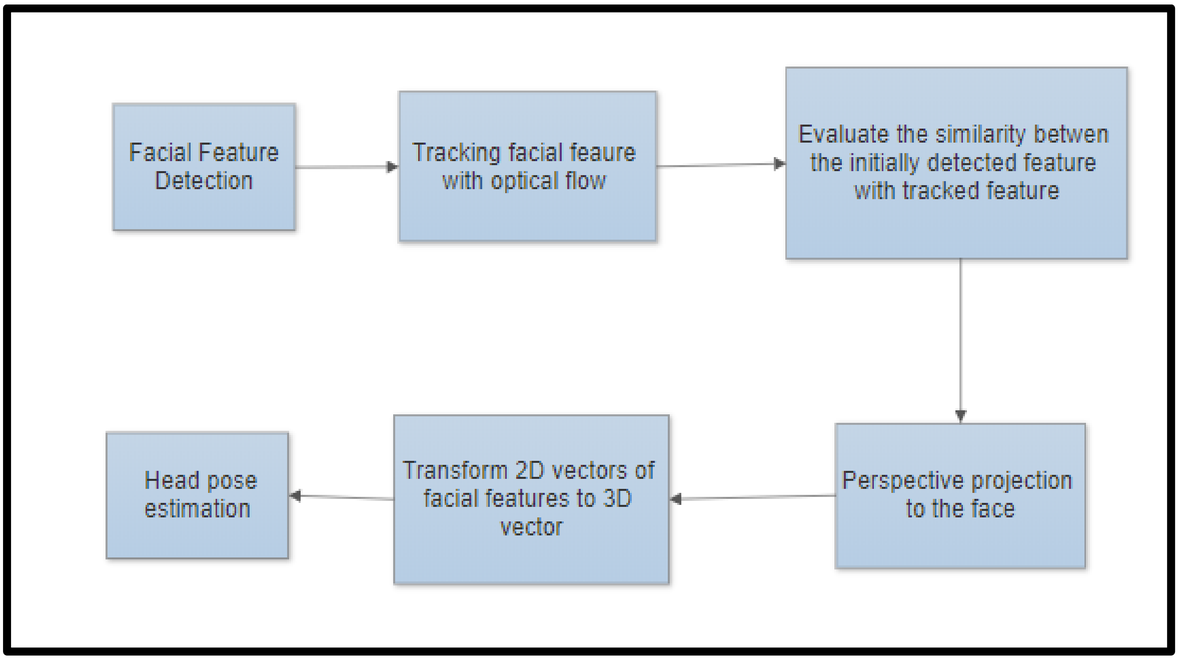

Fig5.3.2(a) Depicts the process of pose estimation.

5.3.3 SILHOUETTE FEATURE EXTRACTION MODULE

With the results of the previous section, a good estimation of the shape can be obtained. However, it only takes into account certain landmarks of the face. The landmarks detected are kept fixed in order to increase stability of the refinement stage. The poses estimated, in the other hand, are used as initialization only and are also further refined considering both landmarks and silhouette information.

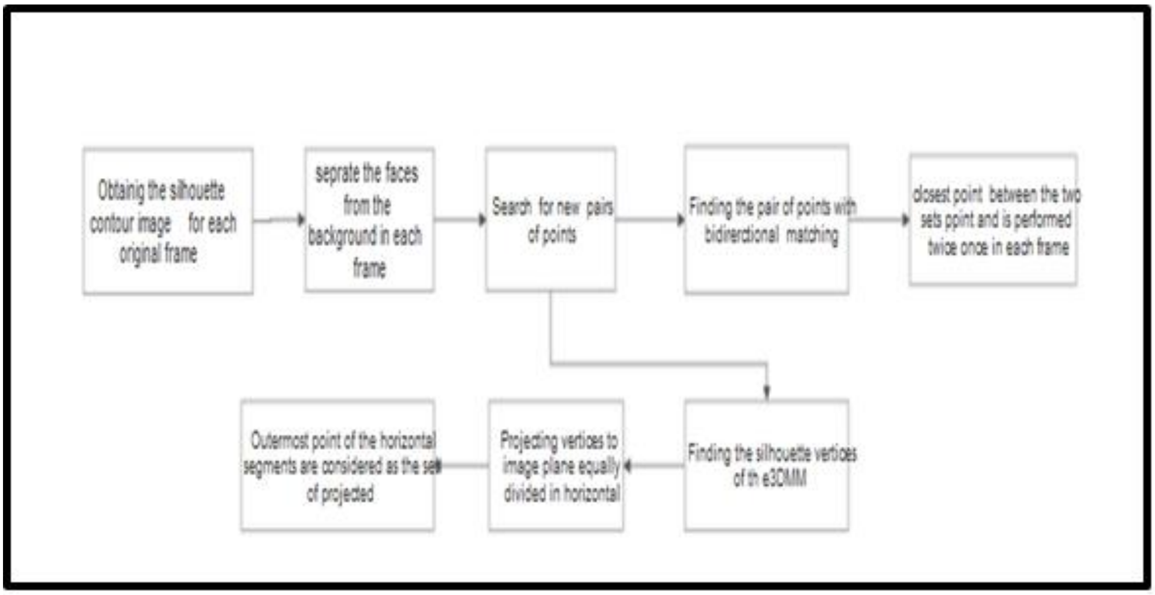

Fig.5.3.3(a) Depicts the process of silhouette feature extraction.

The next section talks about the datasets used.

6. DATASETS USED

This chapter describes about the dataset that has been used to generate 3D face.

Data Courtesy: AFLW2000-3D

AFLW2000-3D is constructed by to evaluate 3D face alignment on challenging unconstrained images. This database contains the first 2000 images from AFLW and expands its annotations with fitted 3DMM parameters and 68 3D landmarks. This database has been used to evaluate the performance of the method used on both face reconstruction and face alignment tasks.

7. IMPLEMENTATION

This chapter describes about the methodology used for the implementation of each module. In mathematics and computer science, an algorithm is an unambiguous specification of how to solve a class of problems. Algorithms can perform calculation, data processing, automated reasoning, and other tasks. This section also describes the framework and the details of our proposed method.

7.1 PROPOSED METHODOLOGY: PRNet

Having a clear description of the methods that will be used to accomplish the project objectives will make a strong application even more competitive. It is the component in the proposal narrative to bridge the gap between the objectives and the eventual outcome. It is also to demonstrate the project's feasibility by detailing the experiences and resources that will be drawn upon to carry out the project.

Firstly, we introduce the characteristics of the position map for our 3D face representation. Then we elaborate the CNN architecture and the loss function designed especially for learning the mapping from unconstrained RGB image to its 3D structure.

We propose a straightforward method that simultaneously reconstructs the 3D facial structure and provides dense alignment. To achieve this, we design a 2D representation called UV position map which records the 3D shape of a complete face in UV space, then train a simple Convolutional Neural Network to regress it from a single 2D image. We also integrate a weight mask into the loss function during training to improve the performance of the network. Our method does not rely on any prior face model, and can reconstruct full facial geometry along with semantic meaning.

Position Map Regression Networks (PRN) is a method to jointly regress dense alignment and 3D face shape in an end-to-end manner.

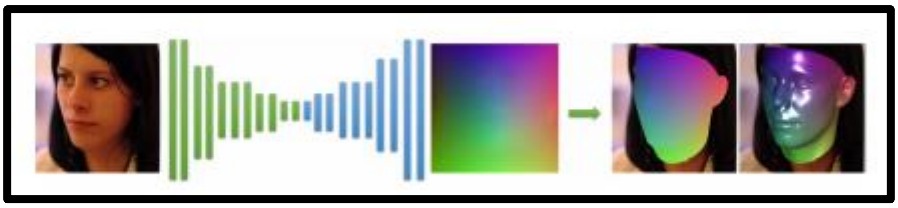

Fig.7.1(a) Depicts the PRN architecture.

7.2 MODULES USED

Large projects are created from multiple small parts that can be specific to a project or can be shared between projects. A module can be a file or a group of files and is called: package, module or component in different languages. The modules can be materialized in source code and can be compiled as shared library or embedded into the main program. A library can be shared between multiple programs.

7.2.1 CASCADE REGRESSION TREE (LANDMARK DETECTION AND POSE ESTIMATION)

Cascaded regression has been recently applied to reconstruct 3D faces from single 2D images directly in shape space, and has achieved state-of-the-art performance. This can be integrated with standalone automated landmark detection methods and reconstruct 3D face shapes that have the same pose and expression as the input face images, rather than normalized pose and expression. The results not only deepen the understanding of cascaded regression-based 3D face reconstruction approaches, but also prove the effectiveness of the proposed method.

Have training data $\{(I_{\pi_i}, \hat{\mathbf{S}}_i^{(t)}, \Delta\mathbf{S}_i^{(t)})\}_{i=1}^{N}$ and the learning rate (shrinkage factor) $0 < \nu < 1$

1. Initialise

$$f_0(I, \hat{\mathbf{S}}^{(t)}) = \arg\min_{\gamma \in \mathbb{R}^{2p}} \sum_{i=1}^{N} \|\Delta\mathbf{S}_i^{(t)} - \gamma\|^2$$

2. for $k = 1, \ldots, K$:

   (a) Set for $i = 1, \ldots, N$

$$\mathbf{r}_{ik} = \Delta\mathbf{S}_i^{(t)} - f_{k-1}(I_{\pi_i}, \hat{\mathbf{S}}_i^{(t)})$$

   (b) Fit a regression tree to the targets $\mathbf{r}_{ik}$ giving a weak regression function $g_k(I, \hat{\mathbf{S}}^{(t)})$.

   (c) Update

$$f_k(I, \hat{\mathbf{S}}^{(t)}) = f_{k-1}(I, \hat{\mathbf{S}}^{(t)}) + \nu\, g_k(I, \hat{\mathbf{S}}^{(t)})$$

3. Output $r_t(I, \hat{\mathbf{S}}^{(t)}) = f_K(I, \hat{\mathbf{S}}^{(t)})$

In the next chapter, we are going to talk about results and discussion.

8. RESULTS AND DISCUSSION

Fig.8(a) Depicts the facial image of a person that has been used for 3D face reconstruction of the same that has been depicted in further image.

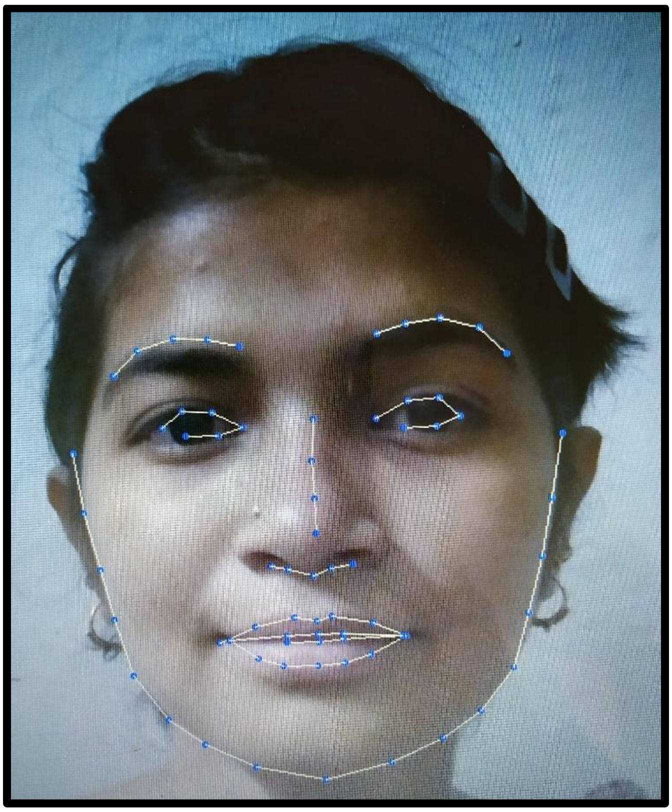

Fig.8(b) Depicts landmark feature extraction.

Fig.8(c) Depicts pose estimation.

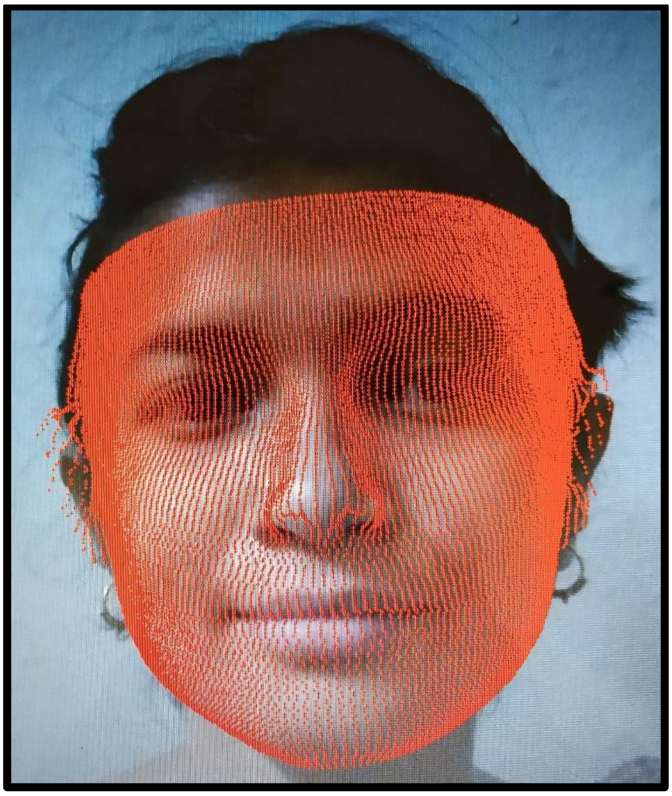

Fig.8(d) Depicts mesh formation.

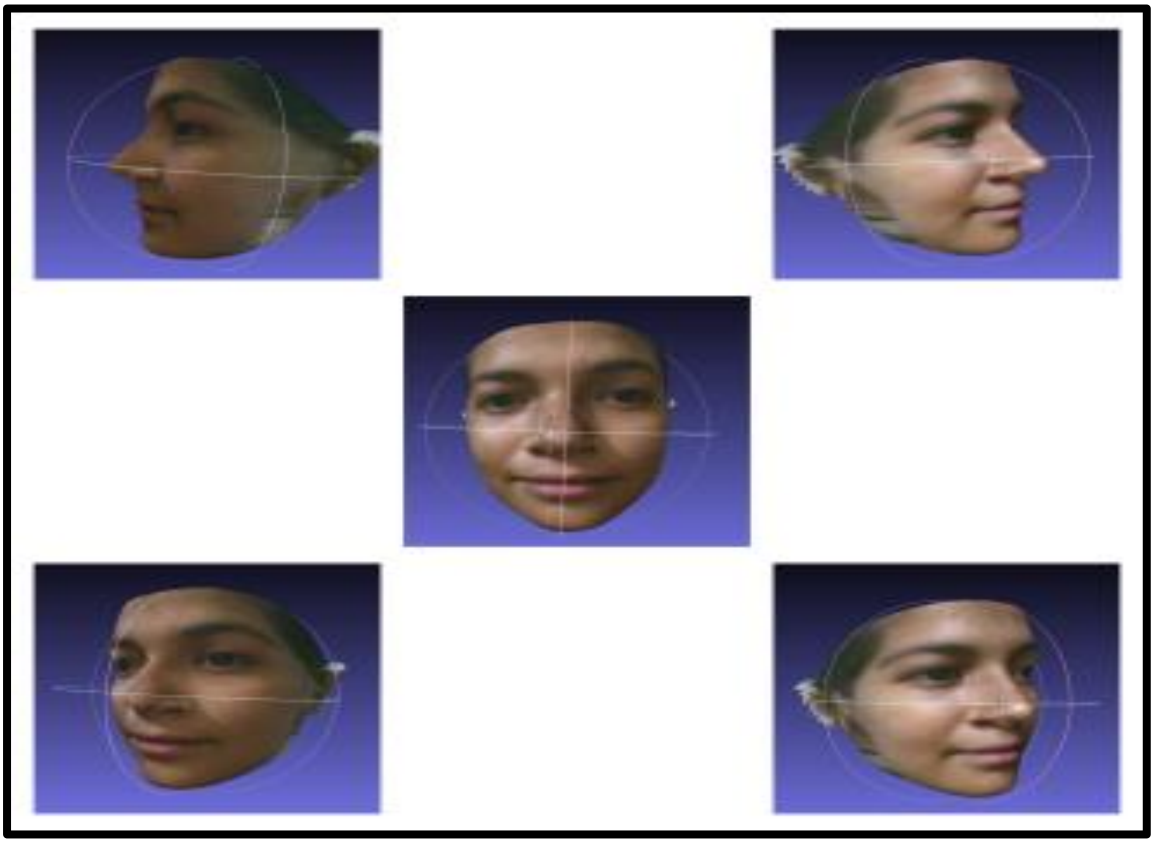

Fig.8(e) Depicts the 3D face reconstruction process and generating various orientations.

9. CONCLUSION

We propose an end-to-end method, which well solves the problems of 3D face alignment and 3D face reconstruction simultaneously. By learning the position map, we directly regress the complete 3D structure along with semantic meaning from a single image. Quantitative and qualitative results demonstrate our method is robust to poses, illuminations and occlusions. Experiments on three test datasets show that our method achieves significant improvements over others. The conclusions are: (i) more landmarks are generally helpful for accurate 3D face reconstruction, but different facial components have different gains from the increased landmarks; (ii) the overall 3D face reconstruction accuracy will be degraded if more areas are covered by the reconstructed 3D faces while the used landmarks remain the same; (iii) the reconstruction accuracy for a specific face area is not affected by the 3D point cloud density in that area or the 3D vertices outside that area as long as the input landmarks are not changed; (iv) using standalone automated facial landmark detection methods together with the cascaded regression based 3D face reconstruction methods is feasible, and the reconstruction accuracy can be improved by disturbing the detected landmarks during training; (v) the cascaded regression based 3D face reconstruction methods have good convergence property.

10. REFERENCES/BIBLIOGRAPHY

[1] Aaron S. Jackson1 Adrian Bulat, "Large Pose 3D Face Reconstruction from a Single Image via Direct Volumetric CNN Regression."

[2] Volker Blanz Thomas Vetter, "A Morphable Model for the Synthesis of 3D Faces."

[3] Ira Kemelmacher-Shlizerman, Ronen Basri, "3D Face Reconstruction from a Single Image Using a Single Reference Face Shape."

[4] Supasorn Suwajanakorn, Ira Kemelmacher-Shlizerman, and Steven M. Seitz, "Total Moving Face Reconstruction"

[5] Pengfei Dou, Shishir K. Shah, and Ioannis A. Kakadiaris, "End-to-end 3D face reconstruction with deep neural networks."

[6] Chuan-Kai Yang · Wei-Ting Chiang, "An interactive facial expression generation system"

[7] Daw-Tung Lin and Han Huang, "Facial Expression Morphing and Animation with LocalWarping Methods."

[8] A neural Algorithm of Artistic Style: Leon A. Gatys, Alexander S. Ecker, Matthias Bethge

Others:

- http://openaccess.thecvf.com/content_ICCV_2017/papers/Jackson_Large_Pose_3D_ICCV_2017_paper.pdf
- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4900890/
- https://arxiv.org/pdf/1604.00449.pdf
- http://cvlab.cse.msu.edu/pdfs/Liu_Zeng_Zhao_Liu_ECCV2016.pdf
- https://www.cv-foundation.org/openaccess/content_cvpr_2015/papers/Kar_Category-Specific_Object_Reconstruction_2015_CVPR_paper.pdf
- http://web.stanford.edu/~zollhoef/papers/EG18_FaceSTAR/paper.pdf
- https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8376005