International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-2, July 2019

# 3D Face Reconstruction Techniques: Passive Methods

**Sincy John, Ajit Danti**

***Abstract:** In the recent literature, 3D face reconstruction received wide interest and has become one of the significant areas of research. 3D face reconstruction provides in depth details on geometrics, texture and color of the face, which are utilized in different applications. It supports a multitude of applications, ranging from face recognition and surveillance to medical imaging, gaming, animation, and virtual reality. This paper attempts to consolidate the research works that have happened in the history of 3D face reconstruction. Also, we try to classify the existing methods based on the input for the process. The databases used in the recent works are discussed and the performance evaluation of methods on different databases is analyzed. The challenges addressed in recent studies are mainly focused on the faster reconstruction of 3D Images, improved accuracy of reconstructed images, human pose identification, image reproduction with higher resolution. Researchers have also tried to address occlusion related problems. Passive methods, used by different researchers are analyzed and their effects on different parameters are discussed in this work. Finally, possible future areas for improvement in terms of reconstructions are presented for the benefit of researchers.*

***Index Terms:** 3D Face modeling, 3D face Reconstruction, computer vision, machine learning, Image Processing.*

## I. INTRODUCTION

Advances in technology have become a boon to numerous applications. As technology is advancing many of the manual processes are being automated. Computer vision and machine learning are two major domains attracting research attention in automation of processes involving visual capabilities. Computer vision has different application areas like authentication and security, health care, industry [1], which considers visual data as inputs and presents decision making as output. The Face is a unique part of the body that is mainly used for the identification process. But it is most complex part of the body with different expressions at each instant of time. Even facial identification confronts with many challenges like, ageing which results in change in face to some extent. Hence, identification based on the face has become a challenging research area. With the advent of 3D technology and computer vision, scope of reconstruction of 3D face is explored in domains like medical imaging, forensics, recognition, animation, virtual reality, movies and telecommunication [2-7]. Compared to a 2D image, object in its 3D model conveys huge amount of information like depth, contours, and views from different perspectives. Humans perceive any object or surrounding as a 3D model due to brain and its complex optic nerve system. Representation of face in three dimensions gives a better understanding of overall face geometry and texture information, which cannot be inferred from a 2D image [7]. 3D face reconstruction can be achieved in two different ways, namely active and passive methods [8-10]. In case of active method, the person interacts with sensors to obtain the 3D representation, whereas in Passive methods 2D images are used. 3D face reconstruction is achieved by generating a three-dimensional image of the face from different projections or an individual 2D image. The reconstruction of images directly from the object using specially made optical systems is often costlier. But, they provide better accuracy than 3D modeling from 2D digital images using machine learning and computer vision algorithms. Hermann Welcker in 1883 and Wilhelm in1895 were the first to recreate 3D facial approximations from cranial remnants. The initial reconstruction technique was clay reconstruction based on approximation [11]. Later on, mathematical models were introduced. Now fully automated reconstruction based approaches based on computer vision, computer geometry and machine learning algorithm are in trend.

### *A. Application Areas*

3D reconstruction has a surfeit of application as it gives significant information based on the area in which it is used. Reconstructed face provides useful information about the different characteristics of face. These include facial texture, colour, size, etc. It provides a high- resolution model with contrast, intensity enhancement and removal of external noise from the 2D image. The application area has been detailed below:

**Forensic**: In forensics, it is mainly used for reconstruction of face from cranial remains of dead bodies. This method is utilized in anthropology to estimate facial structure of primitive men. Modeling faces of culprits from 2D images or sketches are yet another important application in this field. The major problem encountered in this area is the reconstruction of occluded (covered or obstructed by hair, spectacles, cloth, hand gestures) faces. Reconstruction of faces helps when the facial images obtained from the CCTV footage is unrecognizable [6].

**Revised Manuscript Received on July 27, 2019.**

**Sincy John**, Department of Computer Science & Engineering, Faculty of Engineering, Christ (Deemed to be University), Bangalore, India.

**Ajit Danti**, Department of Computer Science & Engineering, Faculty of Engineering, Christ (Deemed to be University), Bangalore, India.

**Medical field**: Mostly used for pre-surgery planning in plastic surgery [2]. It helps the patient and the doctor to have a better understanding of surgery outcome. In the case of reconstructive surgery, portion of the face is reconstructed after an accident or burn. In cosmetic surgery, the 3D face modeling has been done using the modeling software, to provide the patient with better understanding of possible facial feature changes.

**Multimedia and Entertainment Industry**: It is an area where 3D facial images are widely used. Interactive virtual reality system is used to identify human poses, expression and actions using 3D imaging technology and interact with a computer system [3], [4], [5].

**Biometrics**: Facial recognition is used in Security systems and Access control systems in highly secured areas. Very common application is the face unlock in mobiles and laptops [7], [12], [13].

This paper has been structured as follows: The section 2.0 explains steps followed in 3D face reconstruction. The section 3.0 demonstrates on existing reconstruction methods and its advantages and disadvantages, section 4.0 elaborates about the past research in 3D reconstruction techniques and to end with the conclusion and future scope is discussed in section 5.0. Applications areas can be extensively classified as depicted in Fig.1.

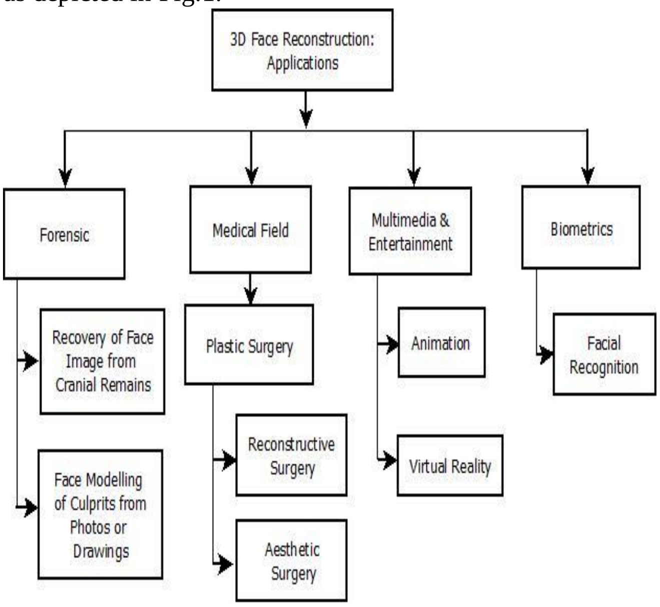


**Fig.1: Applications of 3D Face Reconstruction**

## II. STEPS IN 3D FACE RECONSTRUCTION

Reconstruction of 3D face comprises of different steps. The steps mainly depend on the methods in which reconstruction process is done. The general steps included in the reconstruction process are noise removal, enhancement, facial landmark extraction and 3D face reconstruction. Most of the research works are focused on two categories; facial point detection and the reconstruction techniques. Incremental developments have happened in both the categories, the various techniques used in face reconstruction techniques are explained in the coming sections. The general flowchart of the reconstruction of 3D face is explained in Fig.2.

## III. EXISTING METHODS

Methods used in reconstruction of 3D face are broadly classified into passive and active methods [9][10]. The various techniques can be summarized as follows:

***Active Methods***: These methods rely on the interaction of an energy wave with the object to be reconstructed or the mechanical interference of the object with the system. An energy waves like visible light, laser, ultrasound or microwaves are projected to the object and the reflected wave change in phases and time of flight of the wave to reach back to the source is measured to construct a 3D model. Since the parameters used can be affected by external factors such as exposure, scattering and absorption of the medium [14], a controlled environment is often preferred.

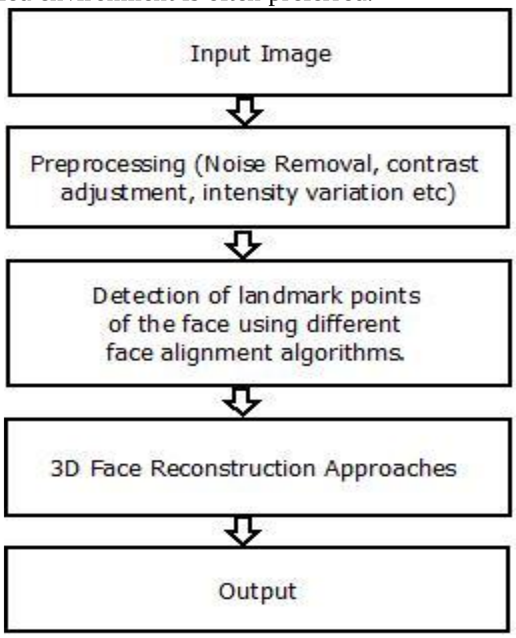


**Fig.2: Typical flowchart of 3D Face reconstruction**

Moreover, most of the 3D modeling systems using active methods consists of specialized and costly hardware equipment that are not easily accessible. These factors limit the application of system modeling leading to implementation of hardware systems specifically made for an application.

***Passive Methods***: These methods do not actively interfere with the target object to construct the model. The reflectance of the surface is captured using a sensor and that information, is used to arrive at the depth, structure and texture of the object. An easily available imaging device like a digital camera may be utilized to capture the 2D image and the 3D model is inferred from the image. Depending on the underlying reconstruction algorithm used, the input can be single image or a sequence of images. These systems are comparatively less expensive and find its applications in a wider domain. Table 1 has depicted the pros and cons of passive and active methods.

## IV. 3D FACE RECONSTRUCTION TECHNIQUES

Reconstruction 3D face is a challenging research area due to many reasons. Firstly, change in facial expression of the individual, secondly the pose of head and neck, thirdly environment factors that affect the images and lastly the occlusion problem that is covering of the face with some obstacles. In the earlier stages, the reconstruction of 3D faces was mainly focused on neutral faces and the images taken in controlled environment [16]. Recent works are concentrated on creating accurate results with wild images and different environment conditions. To tackle these complexities involved in the process, an efficient and a robust technique is needed.

**Table.1; Comparison of active and passive methods**

| Categories | Overview | Pros | Cons |
|---|---|---|---|
| Active Methods (Hardware based Approaches) | 3D faces reconstruction aided by dedicated hardware like radio waves [2] 3D laser Scanners[3][15] | Capable of generating accurate and pragmatic 3D faces | • Leads to additional expenses.<br>• The experimental setup cannot be reproduced easily.<br>• Controlled environment.<br>• Exposure to radiation.<br>• Limiting the usage to certain applications |
| Passive Methods (Software based approaches) | 3D face models are reconstructed from using single image and image sequences from videos. | Used for all the applicatio ns | • 3D face is less realistic compared to active methods.<br>• Occlusion Problem. |

Parke's [17] work related to reconstruction of 3D face was the first attempt to generate a computer model from a 2D image. Henceforth, there had been numerous works published on representation of facial geometry in 3D space with the aid of technology. In [17] Image acquisition was done through two orthogonal photographs of the face marked with polygons. The vertices of these polygons were found out by measuring the distances in the photograph. Further in 1984 Parke's [18] suggested an approach for parameterized representation of human faces as polygonal surfaces. Manipulation of the facial expression through operation like interpolation, rotation, scaling and translation of polygonal surfaces where discussed in the work. A statistical approach was proposed in [19] which discusses about the three dimensional facial model and synthesis of faces. The 3D facial model is created as weighted sum of image bases. The method consisted of two parts; shape and texture synthesis. Former was achieved by summation of weighted shape models and latter by summation of pixel brightness. The decomposition of the procedure allowed analysis and synthesis of face model.

Later a 3D facial model was implemented using frontal and side view image in [20]. While the previous works were aimed at generating generic model, this work addressed the problem of creating a specific 3D model of an individual. The system uses feature extraction, and a nonspecific face model to model the shape of the input face and then applies texture mapping. However, the system was not robust and was susceptible to occlusions, facial features like heavy moustache. However, the output of these techniques was far from a realistic face model.

### *A. Active methods.*

With the advent of depth cameras, and devices like Cyberware, Kinect, 3D laser scanners, researchers were able to approach the problem of reconstruction of three dimensional subjects from a different perceptive. The increased precision in depth information achieved through these devices enabled alternative solutions to some of the existing image reconstruction problems.

The reconstruction of 3D face was done through active methods using Cyberware scanners in [3,15,21]. Cyberware scanners utilize laser scanning method for acquiring the image information. The projected laser light is reflected from scan head to the camera and triangulation method is used to obtain the 3D model [22,23].

Using laser scanned range and reflectance data by a Cyberware scanner in [3], an active method for 3D face reconstruction was proposed. Face mesh is constructed from the range and reflectance data. Relative positions of the facial features, facial muscle convergence and attachment points are identified in the face mesh. An algorithm is used to convert the facial mesh into a simulation of target face capable of creating face expressions. In 1998 Vetter and Blanz [21] suggested a technique in which 3D face model reconstruction is done from a 2D image by approximation method. In this a 3D dataset captured using Cyberware is employed to approximate an adaptable face model and then this generated face is adapted to match the input facial image. Blanz and T. Vetter [15] in 1999 proposed an active method for creating 3D morphable model of human faces using extraneous optical sources such as a laser beam. A 3D morphable model is constructed with the prior knowledge obtained from database constructed using Cyberware using PCA (principal component analysis).Fitting algorithm is used for modifying the 3D model as per the input image. Major disadvantage is that the, 3D morphable model completely depends on the dataset. Performance evaluation of most of the works discussed above was done through visual inspection only, unlike scientific performance evaluation methods used now a days. In medical field 3D facial modeling technique was used for medical imaging like pre-surgery planning and plastic surgery. In 2004 Takács et al [2] designed hardware based face modeling pipeline consisting of, an MRI segmentation module, a 3D surface modeler and a real time engine. In [5] the author has put forward hardware based approach; object is illuminated using structured light projectors and the image is captured using a synchronized video camera. The approach was intended to capture the shape and time varying behaviors.

Photometric stereo is used in [24] is yet another active method for reconstructing 3D face by utilizing near infrared and visible light. Algorithm is proposed to reduce shadow effect by selecting best possible light source for reconstruction. In [25] input image is captured with a TOF (Time of Flight) camera to get depth information used to reconstruct the body pose. A customized Dijkstra's algorithm is used for feature extraction. The local and global pose estimated by combining Hausdorff distance is used to create a late-fusion scheme. Another hardware that is used is the Kinect which has a IR emitter, RGB camera and IR depth sensor [4,26,27]. Figure 3 represents the hardware structure [26]. Zollhofer et al [4] designed a method without human intervention for reconstructing rational 3D face by using depth sensor called Microsoft Kinect. Kinect camera was used to implement 3D colored face model [26] then openCV cascade classifier is used to retrieve the facial feature information from the captured image and kinect fusion algorithm is used for further reconstruction. The hardware employed for 3D face modeling in [27] is kinect2 which gives more precise and accurate details.

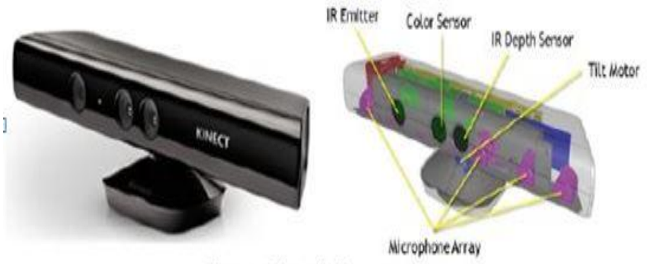


**Fig.3: Microsoft Kinect Camera**

The hardware used is a hand-held device containing structured light is explained in [28] for three dimensional face reconstruction. This scanner is developed in Philips research utilizing smart sensors on an embedded platform. Then a series of steps are performed to obtain the 3D face. [29] lighting calibrations is done on the proxy model obtained from the input image to get a more refined model with fine details. These calibrations simplify the process of obtaining even and hairy regions to reconstruct the 3D face.

### B. *Passive Methods*

This paper explores different techniques under passive methods. There have been a plenty of studies related to the 3D face reconstruction techniques using software methods or passive methods. Fig 4 gives a brief idea about different methods that comes under the passive methods. Some of the relevant literature is discussed in the following paragraphs.

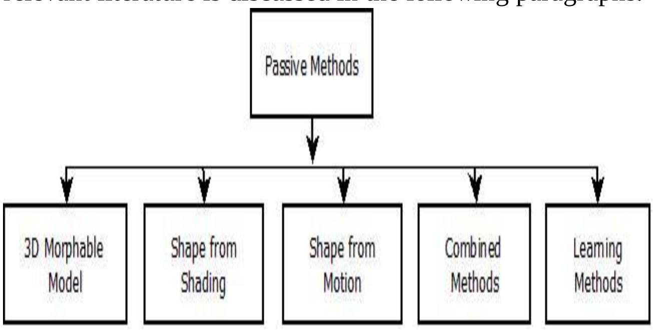


**Fig.4: Passive Methods**

### 3D Morphable Model

In passive methods, reconstruction based on 3D morphable models employing generic or morphable facial models comes under the three dimensional morphable model. One of the breakthroughs in the 3D morphable model used in 3D face reconstruction was in 1999 Blanz and T. Vetter [15] created database of 3D morphable models. Later in 2003 [12], the same authors proposed a model in which the 3D shape reconstruction is done by fitting the three dimensional morphable model to the input image. At first, the common properties of face are studied. Texture and shape parameters are estimated along with 3D parameters. Last is for the recognition purpose faces are represented and compared. A disadvantage was that the 3D morphable model confined to a particular age group and with neutral facial expressions.

Nathan Faggian et al in [30], highlights a technique where a morphable face model is fitted to a sequence of retrieved facial features from multiple views. The 3D morphable model is the progression of the 2D Active Appearance Model (AAM). In different illumination and pose variation, it provides accurate modeling. The author has highlighted about the 3DMM Fitting Algorithms based on feature and texture and in detail explained about the Fitting Morphable Models to 2-Views and n-views. Xiaowen Wang proposed a method to fit a morphable face model to a sequence of extracted features in 2010 [31]. 3D Morphable face model construction is done by computing dense point to point correspondence between the example and reference face. After that the morphable model fitting using shape update algorithm is performed. The advantage is that it can handle the occlusion problem to some extent. The limitation is that prior information about each view is needed.

In 2011, Ira Kemelmacher-Shlizerman and Ronen Basri [32] designed a technique for reconstructing 3D face from a 2D image by the aid of a single 3D face reference model of either a generic face or different individual is used. The advantage of this approach is it avoids the need for a large face database for model creation and robust to unrestrained lighting condition. The disadvantage is performance may degenerate under ambient lighting conditions. The author has briefly explained about the Recovery of Lighting Coefficients, depth recovery, Boundary Conditions for Depth Recovery and Estimating Albedo. Later in 2012 [33], the authors focused in reconstruction of the 3D face by a simplified 3D morphable model (S3DMM) robust to self occlusion. The main motivation for introducing the S3DMM was that the computational cost of the 3D morphable model is high. Here apart from other works cylindrical head model was used, from the know feature points, the occluded points are estimated by 3D model fitting. The main advantage is the low cost and the disadvantage is susceptible to self occlusions caused by head rotation. The disadvantage is that performance totally depends upon the database.

Retrieval Number: B3152078219/19©BEIESP
DOI: 10.35940/ijrte.B3152.078219



Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

**Shape from Shading**

Shape from Shading comes under passive methods; it is the reconstruction of the face from the intensity variations. William A.P. Smith and Edwin R. Hancock in 2006[34], proposed a method for embedding a statistical model of face shape on to a Shape-from-Shading algorithm is presented. This model is trained using a surface normal data acquired from range images and fitted to intensity images using constraints on the surface normal direction as per Lambert's Law. The method makes use of azimuthal equidistant projection, which preserves the distance between the points on the surface of a sphere. This property of the projection is used to calculate a local representation of surface normal's. The author succeeds to demonstrate an efficient and accurate approach using a combination of global statistical constraint and local irradiance constraint. The disadvantage of the method is that only face images with frontal pose can be estimated using this method. In 2011[35], the author has automatically reconstructed 3D face model that is locally consistent with the wild image set. The author has highlighted the different steps involved in the 3D face reconstruction it includes: pose normalization, Initial lighting and shape estimation, Local surface normal estimation, Ambiguity recovery, Integration and Iterations of the algorithm. In this the advantage is that the author has used a wild data collection of varying facial expression, pose age etc. Also as put forth a fully automatic 3D face reconstruction. But the disadvantage was that reconstructed models are not metrically correct and occlusion problems are also encountered.

**Shape from Motion**

In passive method, the next technique is shape from motion in which reconstruction of 3D face is estimated from continuous frames from a single source. Lorenzo Torresani et al proposed reconstruction of time-varying shape and motion of nonrigid 3D objects from 2D using shape from motion (SFM) in 2008 [36]. By using SFM the input video sequence is analyzed to estimate 3D shape. In this Probabilistic Principal Components Analysis (PPCA) is utilized to learn a model of facial deformation. In this texture information is not included just the outline of object is reconstructed.

In 2010 Sung Joo Lee et al was motivated by one of the major issues while using Shape from motion (SFM), the issue is that the reconstructed 3D face is susceptible to point correspondence error due to self occlusion [37]. In order to overcome this issue, the author has tried to overcome this problem by utilizing shape conversion matrix along with shape from shading method to reconstruct the three dimensional face from multiple images. For future work reconstruction of 3D face from non frontal images and face recognition across posses using this approach is focused.

**Combined Methods**

The combination of the above methods ie; shape from shading and 3D morphable model are used in [38]. In this approach, the shape from shading method is used to obtain the surface orientation information which is used has an aid to the fitting phase of 3D morphable model. The major benefit provided by this approach is that a refined 3D morphable model is obtained by using the information obtained from shape from shading method. This indeed has increased the overall accuracy of the reconstructed face. The research gap is to reconstruct realistic face models and to improve the performance of the model fitting step. Occlusion problem is not addressed. J Jo et al designed a combination of shape from motion (SFM) and simplified 3D morphable model method by overcoming the disadvantages of the individual methods to produce better results [39] in 2014. Main advantage of this paper was that this method is robust to pose variation and is captures the attributes of an individual and not influenced by a mean face. The future works are focused to two to three-dimension conversion in three dimensional content services. In this approach the reconstruction of 3D facial shape is done by Shape from Motion using mirrored side and frontal face image generated by bilateral symmetry and S3DMM. J Jo et al [40], proposed feature extraction method where he had considered both the macro and micro level features for constructing the three dimensional face. In this approach combination of Structure from motion, Stereo matching Active appearance model is used. The 3D facial shapes are reconstructed by SFM using the macro- and micro-level FFPs obtained in the two facial images. Benefit of the technique is that it is accurate and person specific with regard to 3D facial image. H.Jin et al presents the reconstruction of 3D faces from images obtained from video sequences [13]. Authors have mentioned about the different approaches used for reconstruction. In this paper a new approach prior constrained structure from motion (PCSFM) is proposed by combining shape from motion and 3D morphable model method. The limitation of this paper is reconstructed face is less realistic and head portions covered by hair that head portion is not reconstructed. The reconstructed face as not included the texture details. Luo Jiang et al [41], reconstructed 3D face by using a coarse to fine optimization strategy. The procedure combines shape from shading and three dimensional morphable model methods.The method utilizes the benefits of bilinear face modeling and SFS, the coarse face-model provides more reliable illumination information thus improving performance of SFS. The SFS based fine modeling provides detailed geometric features, which cannot be provided by coarse face modeling.

**Learning Methods**

In these approaches machine learning approaches are used along with other methods. A learning based approach [42] is used to reconstruct 3D face from a single image. The framework consists of two phases; C-RBF network learning phase and 3D face reconstruction phase. The disadvantage of this reconstruction is that it has not included the texture details, not realistic reconstruction and also occlusion problem is not addressed. The future scope is to overcome the face alignment, to include images with pose variation and 3D faces are not recovered well when compared with the ground truth.

In 2017 [43], 3D reconstruction method based on deep convolution neural network is proposed. In this the 3DMM parameters are regressed using the deep CNN architecture directly from input images. Reconstructed face is resistant to the change in viewing conditions. Based on the data obtained from the 3D model it gives the data correspondence of face data with the data set used. Then the 3DMM parameters are fine tuned to have the 3D reconstruction closer to the input image. But for faces with facial expression the generated image was neutral and missing data shown in the head portions. In the future work more 3DMM parameters will be taken into consideration. Pengfei Dou et al [44] proposed learning based approach with a deep neural network is used for an end to end 3D face reconstruction in which the neutral 3D facial shape and reconstructing the facial expression are both considered as separate subtasks from a single image. The underlying DNN architecture consists of two components, namely a multi-task loss function which is used to divide the 3D face reconstruction into neutral face and expressive 3D facial shape reconstruction. Then a fusion convolution neural network (CNN) to improve facial expression reconstruction is included. In this approach, the reconstructed face does not include the texture information.

In 2018 [45], Position Map Regression Network is proposed to foresee the intense alignment and 3D reconstruction of the input image. In this a UV position map is used for recording 3D positions of all points in UV space from a 2D image. This gives a fully reconstructed 3D face with all the necessary information related to alignment. After this a position map is formed using the UV coordinates of the face surface and the corresponding 3D position of the face structure. Convolution neural network and loss function are used in the learning stage, so that model is able to map unconstrained RGB image to its 3D structure.

Ruiqi Zhao et al [46] proposed a approach that the 3D shape of rigid and non-rigid object is reconstructed from the 2D landmark obtained from the single image. A feed-forward deep neural network algorithm is used to map the 3D face from the 2D landmarks with small reconstruction errors. Multilayer neural network was utilized to recreate a 3D geometric shape from the 2D landmark. The reconstructed face is silhouettes of the images. In this texture details is not included and Occlusion due to hair is addressed.

Table 2 briefs the different passive reconstruction methods with performance evaluation, database used and the results obtained are given.

**Table 2: Comparison for different reconstruction techniques based on passive methods**

| Reference Number | Method | Database | Performance Evaluation | Result |
|---|---|---|---|---|
| [12] | 3D morphable model | CMU-PIE and FERET database. | Face reconstruction accuracy is given in ratio of hit rate against false alarm rate | Hit rate obtained is 77.5% and 87.9% for CMU-PIE and FERET at 1 percent false alarm rate. |
| [30] | 3D morphable model | USF DARPA Human ID 3D Face Database-75 aligned heads | face reconstruction accuracy is given in terms of mean and standard deviation | best results are obtained when regularization constant is $\eta$=10 the mean and standard deviation obtained for yaw-4.281 ± 1.106, pitch -4.220± 1.127, roll- 4.392 ± 1.138 |
| [31] | 3D morphable model | CMU-PIE database | performance evaluation by visual inspection | Nil |
| [32] | 3D morphable model | USF face database- used 77 face models, YaleB face- used as Input image database | Accuracy of the process measured by mean and standard deviation with respect to ground truth | 13.8 ± 2.8mm. |

| | | | | |
|---|---|---|---|---|
| [33] | 3D morphable model | In house Face database with 86 male and 64 female 3D face | root mean square data | RMSE obtained with sample data using QR decomposition has shown better results- 2.25 |
| [34] | Shape from Shading | 3DFS and Yale-B databases. | Based on Illumination, Minimum number of images, Albedo, Reflectance, Alignment, Satisfaction of Data closeness, Approximate shape recovery time | comparison with the work of blanz and vetter in various parameters |
| [35] | Shape from Shading | Labeled Faces in Wild (LFW) | Reconstruction error | For frontal images reconstruction error is 0.76 ±0.48% and 0.93 ±0.68% for frontal and non-frontal. |
| [36] | Shape from Motion | Carnegie Mellon University (CMU) motion capture database | The error in reconstruction was analyzed as a noise function. | with BCD-LS and XCK initiation extact shape is obtained when measurement noise is zero |
| [37] | Shape from Motion | 150 3D face scans with age variation | Root Mean Square error | The average 3D RMS error of the reconstructed 3D face and the ground-truth was less than 3.5mm when the 2D FFPs were annotated automatically. |
| [38] | Combined Methods | Yale-B database | Error rate(%) | the error rate (%) obtained aganinst various illumination techniques is 1.4 |
| [39] | Combined Methods | 150 3D face scans with age variation from 86 males and 64 females | Root Mean Square error | Average 3D RMSE for the recreated 3D face and ground-truth 1.) when facial feature points are annotated manually < 2.6mm 2.) feature points annoted automatically < 3.5mm |
| [40] | Combined Methods | Bosphorus 3D database | Mean error ± Standard deviation | Mean error ± Standard deviation (mm) for 1400 x 1200 pixels is 5.2468 ± 1.1665, 700 x 600 is 6.3258 ± 1.5855, 350 x 300 is 8.7854 ± 2.5151 |
| [13] | Combined Methods | 3D face database | Mean square error and standand deviation | Achieves better mean square error and standand deviation compared to PCA based method |
| [41] | combined method | FACEWAREHOUSE and BFM2009 database | Root Mean Square error | 3D RMSE=1.75±0.29, 3D RMSE of different pose and expression are also mentioned |

| | | | | |
|---|---|---|---|---|
| [42] | learning methods | BU3D database | Mean square error | performance with different subjects ranges from 3.45% to 8.89% when compared to canonical correlation analysis (CCA) and nonlinear manifold embedding and alignment (NMEA) methods in terms of mean square error |
| [43] | learning methods | not specified | Time consumed for a single image | time taken to reconstruct a single image is 0.198s it is less compared to 3DMM and floe based method |
| [44] | learning methods | FRGC2 database, the BU-3DFE database, UHDB31 database | Mean and standard deviation of RMSE (mm) | Mean and standard deviation of RMSE (mm) of various method ranges from 2.73±0.71 to 4.52±1.11 |
| [45] | learning methods | AFLW2000-3D, AFLW-LFPA and Florence | Normalized mean error | Mean NME is 3.7551 , Run time is 9.8 Milliseconds per Image it is less compared to other methods |
| [46] | learning methods | The database used include; fine-grained 3D car (FG3DCar),the Binghamton-Pittsburgh 4D Spontaneous Expression (BU 3DFE),Human 3.6M, CMU Motion Capture. | Reconstruction error | For human face 3D shape reconstruction error is < .004. |

## V. CHALLENGES

The literature reveals that, 100% reconstruction of an image is almost an impossible task. Researchers are working on different methods to get a high definition 3D model closer to the original subject of the individual with lesser error rate and higher efficiency. Table 3 illustrates the methods used in specific, findings and some of the crucial challenges that the researchers are facing is discussed further.

The major research challenges discussed in the recent works include:

1. Reconstruction of input face images with cast shadows [35].
2. Pose variation synthesis of front view of the face from non frontal faces in CCTV footage [39].
3. Self occlusion [33] [41] [47] and occlusion by external objects.
4. The major challenge in 3D morphable model based method is availability of a database which includes variety of faces including different age group and races, as the resultant accuracy depends on models in the database used [32][41].

**Table 3: Different passive methods with their finding and research gap.**

| Reference Number | Methods used in specific | Findings | Research gap |
|---|---|---|---|
| [47] | Local deep feature alignment using a stacked contractive auto-encoder, Feed forward | Authors improved the performance rate compared to CNN, CNN-MM, LDFA-AE, | 3D sample used for training is generated using FaceGen utility hence the result obtained depends |

| | | | |
|---|---|---|---|
| | deep neural network. (Learning Methods) | LDFA-DAE, by minimization of Reconstruction Error. | on the accuracy of FaceGen. Cannot handle pose variation and occlusions |
| [41] | 3D Morphable Model (3DMM) and Shape-from-shading (SFS) (combined methods) | Researchers are able to reconstruct the fine details of the face. They have improved in terms of Mean and standard deviation of root mean square error for pose and expression compared to previous methods. | Variety of faces including different age group and races has not been studied.<br>Recovery of faces when occlusion occurs has not been addressed |
| [46] | feed-forward deep neural network algorithm. (learning methods) | Author's demonstrated feed-forward deep neural network algorithm to get better results in terms of Mean Reconstruction Errors. | Texture details are not considered |
| [40] | Active appearance model and SFM methods. (combined methods) | Results obtained revealed improved output by using Mean error ± Standard deviation (mm) than the earlier methods. They had also demonstrated detailed experiment by using different pixel sized images. | Authors wish to extend their Proposed method to applications like e-commerce and pre-planning Simulations for surgeries using micro-level feature-based 3D facial reconstruction method |
| [39]. | 3D Morphable Model (3DMM) and the Structure from Motion (SfM) methods. (combined methods) | Based on the experimental analysis, authors put forward an enhanced method, which is robust against different poses when judged against the prior methods. | Face recognition of pose variant images and synthesis of front view of the face from non frontal faces in CCTV footage is challenging for the future work. |
| [33]. | simplified 3D morphable model (S3DMM)<br><br>(3D morphable model ) | Compared to previous methods the authors work have shown notable progress in reconstruction performance for self occlusion problem with visible facial feature points. | Pose variations and problems of Self occlusion to be addressed in further proficient way. |
| [35]. | Shape from shading | First method to automatically generate 3D faces from unconstrained images. | Reconstruct input images with cast shadows is not addressed. |

## VI. CONCLUSION AND FUTURE SCOPE

3D Facial reconstruction provides more comprehensive information about a subject face compared to a 2D image. This paper investigates about the various techniques that are used for 3D face reconstruction; it also gives brief idea about the different application areas and about the broad classification of the existing methods.

Active methods provide more realistic models and it is very much useful for medical application and a dedicated step up has to be maintained. Passive methods work on images and require less sophisticated experimental setup and hence is less expensive. Each paper has its pros and cons. The selection of reconstruction technique varies according to the application objective. Over the years, with the progress in image capturing technology and image processing, the quality and accuracy of 3D face reconstruction has been improved a lot. An area of further scope of exploration is application of different learning techniques to create intelligent systems that gives better reconstruction results for the wild images. This would require the model to be trained on highly versatile database which covers various illumination conditions, pose variations, ethnicity, age groups etc. The idea of hybrid methods can throw light on the challenging areas that the researchers are facing now by taking advantage of different methods to overcome limitations of a particular model. Another aspect under consideration is obtaining high definition 3D face with minimal loss in details like texture, wrinkles, and contours are a real challenge.

## ACKNOWLEDGMENT

We thank Christ (Deemed to be University) for providing the necessary arrangement and facility for carrying the research work.

## REFERENCES


1. G. Sansoni, M. Trebeschi and F. Docchio, "State-of-The-Art and Applications of 3D Imaging Sensors in Industry, Cultural Heritage, Medicine, and Criminal Investigation," *Sensors*, vol. 9, no. 1, pp. 568–601, 2009.
2. B. Takács, S. Pieper, J. Cebral, B. Kiss, B. Benedek, and G. Szijártó, "Facial Modeling for Plastic Surgery Using Magnetic Resonance Imagery and 3D Surface Data," *Three Dimens. Image Capture Appl. VI SPIE Electron. Imaging*, no. January, pp. 18–22, 2004.
3. Y. Lee, D. Terzopoulos, and K. Waters, "Realistic Modeling for Facial Animation," in *SIGGRAPH '95 Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*, vol. 95, pp. 55–62, 1995.
4. M. Zollh, M. Martinek, M. Stamminger, and S. Jochen, "Automatic Reconstruction of Personalized Avatars from 3D Face Scans," *Comput. Animat. Virtual Worlds*, vol. 22, no. 2, pp. 195–202, 2011.
5. L. Zhang, N. Snavely, B. Curless, and S. M. Seitz, "Spacetime Faces : High-Resolution Capture for Modeling and Animation," *In: Deng Z., Neumann U. (eds) Data-Driven 3D Facial Animation. Springer, London*, vol. 23, no. 3, pp. 248–276, 2008.
6. C. van Dam, R. Veldhuis and L. Spreeuwers, "Face reconstruction from image sequences for forensic face comparison," in *IET Biometrics*, vol. 5, no. 2, pp. 140-146, 6 2016.
7. G. Berk, A. A. Salah and L. Akarun, "3D Face Recognition : Technology and Applications 3D Face Recognition" *In: Tistarelli M., Li S.Z., Chellappa R. (eds) Handbook of Remote Biometrics. Advances in Pattern Recognition. Springer, London,* pp. 217-246, 2009.
8. F. Remondino and S. El-Hakim , "Image-Based Modeling : A Review," *Photogramm. Rec.*, vol. 21, no. 115, pp. 269–291, 2006.
9. G. Stylianou and A. Lanitis, "Image Based 3D Face Reconstruction : A Survey," *Int. J. Image Graph.*, vol. 9, no. 2, pp. 217–250, 2009.
10. S. Herbort and C. Wöhler, "An Introduction to Image-based 3D Surface Reconstruction and a Survey of Photometric Stereo Methods," *3D Res.*, vol. 03, pp. 1–17, 2011.
11. J. Omstead, "Facial Reconstruction," *Totem Univ. West. Ontario J. Anthropol.*, vol. 10, no. 1, pp. 37–46, 2002.
12. V. Blanz and T. Vetter, "Face recognition based on fitting a 3D morphable model," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 9, pp. 1063-1074, Sept. 2003.
13. H. Jin, X. Wang, Z. Zhong, and J. Hua, "Robust 3D face modeling and reconstruction from frontal and side images," *Comput. Aided Geom. Des.*, vol. 50, no. January, pp. 1–13, 2017.
14. G. Bianco, A. Gallo, F. Bruno, and M. Muzzupappa, "A Comparative Analysis between Active and Passive Techniques for Underwater 3D Reconstruction of Close-Range Objects," *Sensors*, vol. 13, no. 8, pp. 11007–11031, 2013.
15. T. Vetter, "A Morphable Model For The Synthesis Of 3D Faces," in *SIGGRAPH '99 Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, 1999, pp. 187–194.
16. Y. Wu and Q. Ji, "Facial Landmark Detection : A Literature Survey," *Int. J. Comput. Vis.*, vol. 127, no. 2, pp. 115–142, 2019.
17. F. I. Parke, "Computer generated animation of faces," in *ACM '72 Proceedings of the ACM annual conference - Volume 1*, pp. 451–457, 1972.
18. F. I. Parke, "A Parametric model for human faces," Doctoral Dissertation, The University of Utah 1976.
19. C. S. Choi, T. Okazaki, H. Harashima and T. Takebe, "A system of analyzing and synthesizing facial images," *1991., IEEE International Sympoisum on Circuits and Systems*, Singapore, vol.5, pp. 2665-2668,1991.
20. T. Akimoto, Y. Suenaga and R. S. Wallace, "Automatic creation of 3D facial models," in *IEEE Computer Graphics and Applications*, vol. 13, no. 5, pp. 16-22, Sept. 1993.
21. T. Vetter and V. Blanz, "Estimating coloured 3D face models from single images : An example based approach," in *European conference on computer vision*, Springer, Berlin, Heidelberg, pp. 499–513, 1998.
22. C. L. Istook and S.J. Hwang, "3D body scanning systems with application to the apparel industry," *J. Fash. Mark. Manag. An Int. J.*, vol. 5, no. 2, pp. 120–132, 2001.
23. C. L. Istook, *Three-dimensional body scanning to improve fit*. Woodhead Publishing Limited, 2008.
24. M. F. Hansen, G. A. Atkinson, L. N. Smith, and M. L. Smith, "3D face reconstructions from photometric stereo using near infrared and visible light," *Comput. Vis. Image Underst.*, vol. 114, no. 8, pp. 942–951, 2010.
25. A. Baak, M. Müller, G. Bharaj, H. Seidel and C. Theobalt, "A data-driven approach for real-time full body pose reconstruction from a depth camera," *2011 International Conference on Computer Vision*, Barcelona, pp. 1092-1099, 2011.
26. O. Ziskind, R. Mordoch, N. Toledano, A. M. Sela, and Y. Honen, "Generating 3D Colored Face Model Using a Kinect Camera," 2015.
27. J. Li, Y. Zhang, P. Xu, S. Lan and S. Li, "3D Personalized Face Modeling Based on KINECT2," *2016 9th International Symposium on Computational Intelligence and Design (ISCID)*, Hangzhou, pp. 197-201, 2016.
28. R. Villescas, "3D Face Reconstruction using Structured Light on a Hand-held Device," 2013.
29. X. Cao, Z. Chen, A. Chen, X. Chen, S. Li and J. Yu, "Sparse Photometric 3D Face Reconstruction Guided by Morphable Models," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, pp. 4635-4644, 2018.
30. N. Faggian, A. Paplinski and J. Sherrah, "3D Morphable Model fitting from multiple views," *2008 8th IEEE International Conference on Automatic Face & Gesture Recognition*, Amsterdam, 2008, pp. 1-6.
31. X. Wang, W. Liang and L. Zhang, "Morphable Face Reconstruction with Multiple Views," *2010 Second International Conference on Intelligent Human-Machine Systems and Cybernetics*, Nanjing, Jiangsu, pp. 250-253, 2010.
32. I. Kemelmacher-Shlizerman and R. Basri, "3D Face Reconstruction from a Single Image Using a Single Reference Face Shape," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 2, pp. 394-405, Feb. 2011.
33. Y. J. Lee, S. J. Lee, K. R. Park, J. Jo, and J. Kim, "Single view-based 3D face reconstruction robust to self-occlusion," *EURASIP Journal on Advances in Signal Processing*, no. April, 2012.
34. W. A. P. Smith and E. R. Hancock, "Recovering Facial Shape Using a Statistical Model of Surface Normal Direction," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, pp. 1914-1930, Dec. 2006.
35. I. Kemelmacher-Shlizerman and S. M. Seitz, "Face reconstruction in the wild," *2011 International Conference on Computer Vision*, Barcelona, pp. 1746-1753, 2011.

36. L. Torresani, A. Hertzmann and C. Bregler, "Nonrigid Structure-from-Motion: Estimating Shape and Motion with Hierarchical Priors," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 5, pp. 878-892, May 2008.
37. S. Joo, K. Ryoung, and J. Kim, "A SfM-based 3D face reconstruction method robust to self-occlusion by using a shape conversion matrix," *Pattern Recognition*, vol. 44, no. 7, pp. 1470–1486, 2011.
38. A. Patel and W. A. P. Smith, "Driving 3D morphable models using shading cues," *Pattern Recognition*, vol. 45, no. 5, pp. 1993–2004, 2012.
39. J. Jo, H. Choi, I. Kim, and J. Kim, "Single-view-based 3D facial reconstruction method robust against pose variations," *Pattern Recognition.*, vol. 48, no. 1, pp. 73–85, 2014.
40. J. Jo, H. Kim, and J. Kim, "3D Facial Shape Reconstruction Using Macro- and Micro-level Features from High Resolution Facial Images," *Image Vis. Comput.*, vol. 64, no. August, pp. 1–9, 2017.
41. L. Jiang, J. Zhang, B. Deng, H. Li and L. Liu, "3D Face Reconstruction With Geometry Details From a Single Image," in *IEEE Transactions on Image Processing*, vol. 27, no. 10, pp. 4756-4770, Oct. 2018.
42. M. Song, D. Tao, X. Huang, C. Chen and J. Bu, "Three-Dimensional Face Reconstruction From a Single Image by a Coupled RBF Network," in *IEEE Transactions on Image Processing*, vol. 21, no. 5, pp. 2887-2897, May 2012.
43. L. Fangmin, C. Ke and L. Xinhua, "3D Face Reconstruction Based on Convolutional Neural Network," *2017 10th International Conference on Intelligent Computation Technology and Automation (ICICTA)*, Changsha, pp. 71-74, 2017.
44. P. Dou, S. K. Shah and I. A. Kakadiaris, "End-to-End 3D Face Reconstruction with Deep Neural Networks," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, pp. 1503-1512, 2017.
45. Y. Feng, F. Wu, X. Shao, Y. Wang, and X. Zhou. "Joint 3D face reconstruction and dense alignment with position map regression network". *In CVPR, arXiv* preprint arXiv:1803.07835.
46. R. Zhao, Y. Wang and A. M. Martinez, "A Simple, Fast and Highly-Accurate Algorithm to Recover 3D Shape from 2D Landmarks on a Single Image," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 12, pp. 3059-3066, 1 Dec. 2018.
47. J. Zhang and C. Zhu, "Approach to 3D face reconstruction through local deep feature alignment," in *IET Computer Vision*, vol. 13, no. 2, pp. 213-223, 3 2019.

## AUTHORS PROFILE

**Sincy John** is currently pursuing her PhD in Image Processing from Faculty of Engineering, Christ Deemed to be University, Bengaluru, India. Her research interests include Image Processing, 3D face reconstruction and Computer Vision.

.

**Dr. Ajit Danti** is currently working as Professor in Department of CSE, Faculty of Engineering, Christ Deemed to be University, Bengaluru, India. He received his PhD from Gulbarga University, Gulbarga, India. His research interest includes Image Processing, Face Recognition, Computer Vision