

# Method Comparison of 3D Facial Reconstruction Corresponding to 2D Image

**H P A Tjahyaningtjas<sup>1\*</sup>, P Puspitasari<sup>1</sup>, Y Yamasari<sup>2</sup>, L Anifah<sup>2</sup> and IGP A Buditjahyanto<sup>2</sup>**

<sup>1</sup>Electrical Engineering Department, Faculty of Engineering, Universitas Negeri Surabaya, Surabaya, Indonesia

<sup>2</sup>Informatics Engineering Department, Faculty of Engineering, Universitas Negeri Surabaya, Surabaya, Indonesia

\*hapsaripeni@unesa.ac.id

**Abstract.** Recent years, facial reconstruction becomes a much-studied topic. In particular, the past decade has witnessed a renewed interest, generating a large number of research centers and proposing techniques to address them but most have their limitations and drawbacks. The main constrain of attaining an accuracy of the facial information for reconstruction of its 3D counterpart was increasing accuracy for 3D face geometry. Though this topic has gained a lot of concern and popularity, a fully accurate facial reconstruction mechanism has not yet been identified due to the complexity and ambiguity involved. This survey focuses on 3D face reconstruction by presenting comparison methods between Shape-from-Shading, the 3D Morphable Model and Structure from Movement based on 2D face analysis with higher accuracy. 3D face model density affects the provided information. The dense of the 3D facial model is; the more information it could provide. 3D facial reconstruction method currently requires a complicated process and high system costs.

## 1. Introduction

As the development of technology in the field of Information Technology and Computers, Three-dimensional technology (3D) is also growing rapidly. The need of 3D objects visualisation is widely used in animation and graphics applications, architecture, education and cultural recognition, as well as virtual reality.

Humans can perceive 3D shape just by looking at the 2D image. The human brain plays an important role in the formation of 3D imagery through 2D imagery. After looking at 2D image, the human eye signals to the brain about the objects which is captured through nerve signals. After the nerve signals are processed, the brain forms a 3D image of a 2D object. This usefulness is a human unconscious behavior when simulated using a computer then its effectiveness and accuracy is still a big problem in the way of image review for 3D.

The face is the outer part of the body which is very important because it can be used as a personal identification. Facial images can be found on photos and digital images or 2D facial images. To get a specific 3D model image for the purpose of being used in many applications such as 3D assisted face recognition, 3D expression recognition, and facial animation, researchers are still working to solve the



problem of obtaining 3D images of image formation through the use of 3d sensors or 3d image formation obtained from the 2D image.

One of the problems that arise in the process of facial recognition is the difference in position, lighting, and the expression on each image of the 2D face that is varied. It can be simplified by the existence of 3D face geometry. 3D face geometry can produce facial images with different positions, exposure, and lighting.

In general, 3D face reconstruction is divided into three categories:

**Stereo Vision based methods:** This method performs face construction using optical image technology. These include shapes from shading (SFS) and Binocular Stereo Vision (BSV) methods. Assuming, SFS can predict the 3D face of the gradual variation of the shading in an image [1]. BSV is also aimed at estimating the depth faces from the 2D image directly [2]. It reconstructs the 3D face from a pair of calibrated images.

**Deformation based methods:** The aim of the methods is fitting parametric or fixed 3D face models to the given 2D image or the 3D points which is estimated from the 2D image to reconstruct the 3D face. 3D morphable model (3DMM) is one type of the methods which fits the model to the whole 2D image or full 3D points cloud [3]. The method can automatically generate 3D faces in different poses and illuminations.

**Learning based methods:** The basic idea of the methods is exploring or constructing underlying mapping between 2D and 3D faces in a hidden space rather than in the original spaces. The reason is that it is usually difficult to find the accurate mapping between original 2D and 3D faces, as most of them are located in high dimensional spaces. Hence, the learning based method should explore: 1) mapping between original 2D faces and their intrinsic representations, and mapping between original 3D faces and their intrinsic representations, respectively; and 2) mappings between intrinsic representations of 2D faces and those of 3D faces in the hidden space.

There are several types of the learning based methods. CCA is a common method that is maximizing the correlation between projections of two training data sets to explore their relationship [4]. PLS is another type of learning based methods to overcome this problem [5]. It aims at maximizing covariance between two data sets to find a regression matrix between them. Later, based on a well-known assumption that 2D and 3D faces are located on smooth manifolds, Song [6] proposed a couple radial basis function (C-RBF) network to reconstruct 3D faces. They trained the C-RBF network with a pair of 2D and 3D faces to find their intrinsic representations and to explore the nonlinear mapping between them and then reconstructed the particular faces by using a linear combination of their nearest neighbours.

In addition, the approach which is mentioned above, there are other approaches, including approaches that combine two or more approaches to produce realistic faces.

This paper describes various approaches in 3D face reconstruction of 2D imagery, its step approach and the constraints faced in facial reconstruction. At the end of the paper, it will be concluded some speculation about 3D face reconstruction of the 2D image in the future.

## 2. General approach

There are many approaches for reconstructing 3D faces but the choice of approach may vary according to the application in which the reconstruction is used. There are two standard methodologies which normally received to make human facial 3D models. One of the approaches is to utilize particular 3D capture systems like 3D scanners to capture texture notwithstanding profundity i.e., 3D shape. The high cost and speed constraints of present 3D scanning devices are the conspicuous deficiencies to get adequate and helpful information. The second approach is to create calculations to reproduce 3D face models from a 2D image, for example, video arrangements or multi-see photos. This approach, call 3D reconstruction algorithm is a critical apparatus that can be utilized as a part of observation and in different interactive media applications. Amid the previous decade, numerous 3D recreation calculations have been produced and can be arranged in to four gatherings, shape-from-shading (SFS)

[7,8], the 3D Morphable Model (3DMM)[9,10], structure from movement (SfM)[11,12] and learning [13,14].

The most successful method was a structure from motion, Structure from motion (SfM) is a popular approach to recover the 3D shape of an object when multiple frames of an image sequences are available. In SFM, the 3D information about a collection of discrete structures, such as lines, curves, and points, is recovered from a 2D collection of lines, curves, and points. 2D images are formed by projections from the 3D world. As from the literature two well-known projection models are available, the perspective and the orthographic [15],[16],[17]. Multiple 2D facial images including the frontal face are used to reconstruct the 3D Face [18]. The Bosphorus subject data base is used in this experiment with 24 landmark features. Differential Evolution Optimization based algorithm to estimate the pose and reconstruct the 3D facial shape from 2D face images. The proposed algorithm requires one non-frontal-view and one frontal view face image to reconstruct the sparse 3D face model. The optimization is carried out in two steps. Further simple model integration is proposed for the cases when more than one non-frontal view image is available to improve the accuracy of the 3D model estimation.

Shape from shading (SfS) remains important because it ignores most of the depths of defocus. Algorithms to recover shapes from general shading are generally considered to produce excellent results in global minimization while local approaches are more erroneous but faster [19]. Developed methods for facial shape recovery from a single image that uses a reference 3D face model from a different individual or generic face. Intuitively, this method uses an input image as a guide for "forming" the reference model to achieve the desired reconstruction. Here use shading information to restore 3D face shapes when using reference forms and extract important information for unrecognized a priori recovery processes, such as lighting and poses. Here are the key points to work with 3D face reconstruction. : 1) 3D face reconstruction using 2D images using data mining algorithm; 2) Various approaches have been proposed as a solution to this problem but most have limitations and disadvantages; 3) The shape of the shading, the shape of the silhouette, the shape of movement and the analysis by synthesis using the current morphable model is considered as the primary method for achieving face information for the reconstruction of its 3D counterpart; 4) Although this topic has gained a lot of importance and popularity, the accuracy of face reconstruction mechanisms have not been fully identified due to the complexity and ambiguity that are involved. [20]

The use of the 3D morphable model (3DMM) has been widely applied to face analysis since this model was not affected by intra-personal variations such as pose and illumination to produce an ideal representation of 3D faces. This method can make the construction process with the right shapes and textures, pose and illumination via a fitting process. In addition, Morphable model can model intra-personal variations such as pose, illumination, expression. Given a single facial input image, it can recover both inter-personal (3D shape and texture) and intra-personal properties (pose, illumination) via a fitting algorithm.

### **3. Steps in a regular 3D face reconstruction approach**

After considering all of these approaches, a set of general steps can be derived that will be included in the usual 3D face reconstruction algorithm. The following is a list of the steps identified below:

#### *3.1 Repairing the damaged areas (caused by noise, occlusion or shadows)*

The input image conditions may not always be satisfactory; They may be damaged or corrupted. Noise from the image, if any, may lead to inaccurate reconstruction. Shadows, poor lighting conditions, and occlusion prevent accurate feature extraction on the face. For this reason, these damaged areas must be eliminated prior to reconstruction Face localization.

#### *3.2 Facial component detection*

After the local face is separated, the parts of the face can be effectively recognized. Picture based systems, outlines, and highlight focuses can be utilized to identify these facial parts. In distinguishing

these facial parts, perceiving the two corners of the eyes, the tip of the nose and the middle and the end purposes of the mouth would demonstrate enough.

### 3.3 Depth estimation

For an accurate and realistic reconstruction, both location and depth of the facial features of the reconstructed face should be equivalent to the real face. Constructing the depth map of the input image will assist in depth estimation.

For a precise and practical remaking, both location and depth of the facial features of the reconstructed face ought to be equal to the genuine face. Developing the depth map of the input image will aid in depth estimation.

### 3.4 3D face reconstruction

After face segments' areas and profundity are distinguished the 3D face can be recreated. A default 3D model can be distorted by the genuine elements to acquire the final 3D face. The texture ought to be mapped onto the 3D face. This is a mind boggling process since the texture information is picked up from 2D space which must be mapped onto a 3D space. Some methodologies extend the frontal image straightforwardly onto the 3D face yet in the event that the approach takes multiple input images these can be twisted into the texture space to produce a more reasonable impact.

## 4. Recent work

Baumberger et al using a new approach strategy to detect landmarks and poses based on 3DMM models [21]. The new approach combines sparse data and silhouette information. In his research, he investigates the effect of many input frames containing important landmark information not only to improve the accuracy of reconstruction but also the shape of modeling. The three-step solution which is offered are facial landmarks detection and pose estimation, 3D shape estimation including silhouette information, and finally texture mapping. Xiaoping et al. [22] proposes an efficient framework for automatically reconstructing 3D face shapes and textures. In his research, the model is altered through a set of sparse feature points that are obtained from the reliable landmarks detection algorithm [23]. It is found that the position of the eyes, nose, and mouth are almost positioned invariant, but not on the side of the face. The five steps offer to evaluate the relationship between reconstruction accuracy and the number of different points are: 1) train a morphable model on the 3D database; 2) detect 68 landmarks on the face in the input image; 3) do an analysis-by-synthesis loop to optimize shape through 23 landmarks; 4) extract and infer texture from input image; 5) refine the full 3D reconstruction

Kemelmacher *et al* [24] is Introducing a new method for facial shape recovery from a single image that uses only a single reference 3D face model from a different individual or generic face. This method uses an input image as a guide forming a reference model to achieve the desired reconstruction. Shading information is used to improve the face of the 3D shape, lighting and poses are used to extract important information when the repair process uses the reference shape and albedo. The input image conditions may not always be satisfactory; They may be damaged or damage. The input image's condition, lead to inaccurate reconstruction. Shadows, poor lighting conditions, and occlusion prevent accurate feature extraction on the face. For this reason, these damaged areas must be eliminated prior to reconstruction [15].

**Table 1.** Method for 3d Facial Reconstruction Corresponding to 2D image

Reference	Method	2D used	Landmark	Database	Median error
Christian Baumberger[21]	3D MM and Shillouete	Y	20	Private	1,11
Ira Kemelmacher [24]	SfS	Y	70 6	USF	6,5 $\pm$ 1,4
Shu Zhiang [23]	3DMM and SfM	Y	automatic	Private	1,31
Seema Chavan [20]	SfS	Y	22	Private	2,2
Phunnam Chandar [15]	SfM	Y	24	3D Bosphorus	0,9096
Xiaoping Hu [22]	3DMM	Y	68	BJUT	3,34
Boulbaba Ben Amor [19]	SfM	Y	30	Private	1,62
Ramasubramanian, M [25]	3D MM	Y	29	Private	1,675
J. Fortuna [12]	SfM	Y	50	Private	2,26

Shu Ziang et al produces a new method combining 3DMM and SfM [23]. It can reconstruct a very solid 3D human face. In addition, it only takes one frame of data from an off-the-shelf RGB-D sensor, even low-resolution low-cost sensors. The process is automatic without manual intervention. The proposed method is divided into two main stages: (a) the initial 3D cloud point acquisition and (b) the cloud propagation of a dense face point. The following methods outline below: 1) Extract one frame data from the RGB-D sensor, which contains the RGB image and the Depth image. With sensor parameters, the registration between the two images is done to obtain 2D-3D correspondence between RGB image and a Depth image; 2) Face recognition is done on RGB image. With the face area which is obtained on RGB images and 2D-3D correspondences between RGB and Depth data, the depth image is cut to store data that only belongs to the face area. To further refine the face region in the depth data, the k-means clustering algorithm is applied to eliminate non-face depth data. At the end of this step, a 3D starting point cloud can be obtained with the proper human face region; 3) Approximate Nearest Neighbor (ANN) algorithm is applied to calculate neighboring structures from the cloud point obtained in the previous step. With a neighboring cloud, an interpolation that is based on Radial Basis Function (RBF) is used to deploy a solid 3D dot cloud from the cloud's starting point. By combining 3DMM and SfM methods and applying ANN algorithm for analysing the 3D face, both methods lead the minimum median error (Table 1).

## 5. Conclusion

The 2D picture of a face is exceptionally touchy to changes in head pose and expressions so a fruitful reconstruction approach ought to have the capacity to remove these face details. These Methodologies are based on silhouettes and prior knowledge can be worthwhile in tending to this issue. While joining the two methods could vigorous the accuracy of 3D face reconstruction. While remaking 3D faces from 2D images the key wellspring of data is the intensity based features and landmarks of the image. However, intensity alone is insufficient in case of low intensity, noise, occlusion, illumination variations and/or shadows are being available in the input image. The anatomical landmark is contended to be a more exact wellspring of information, yet they are somewhat thin and hard to find.

Almost all research on facial reconstruction is limited only in the front area only. Facial reconstruction for a wider area including the face with hair and ears should be developed. Topics of 3D face reconstruction research from the 2D image in the field of computer are significant enough to develop, further research related to facial expression analysis and its application in video conference



still need to be developed. The 3D face reconstruction can be extended to produce aging software which has the capability to produce younger or the older face of the input image.

## References

- [1] B K Horn 1989 *Obtaining shape from shading information* MIT press.
- [2] N Uchida, T Shibahara, T Aoki, H Nakajima and K Kobayashi 2005 3D face recognition using passive stereo vision *Computer Vision and Pattern Recognition 2005 CVPR. IEEE Computer Society Conference II*–950
- [3] V Blanz and T Vetter 1999 A morphable model for the synthesis of 3D faces *Proceedings of the 26th annual conference on Computer graphics and interactive techniques* 187–194
- [4] M Reiter, R Dormer, G Langs and H Bischof 2006 3D and infrared face reconstruction from rgb data using canonical correlation analysis *Proc. Pattern Recognition ICPR 2006 18th International Conference* 425–428
- [5] M Castel'an and J Van Horebeek 2008 3D face shape approximation from intensities using partial least squares *Proc. Computer Vision and Pattern Recognition Workshops CVPRW'08 IEEE Computer Society Conference* 1–8
- [6] M Song, D Tao, X Huang, C Chen and J Bu 2012 Three-dimensional face reconstruction from a single image by a coupled RBF network *IEEE Trans. Image Process.* **21** 5 2887–2897
- [7] A Thelen, S Frey, S Hirsch and P Hering 2009 Improvements in Shape-from-Focus for Holographic Reconstructions with Regard to Focus Operators, Neighborhood-size, and Height Value Interpolation *IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society* **18** 151–157
- [8] M Castel and E R Hancock 2006 Acquiring Height Data from a Single Image of a Face using Local Shape Indicators *Computer Vision and Image Understanding* **103** 1 64–79
- [9] D Jiang, Y Hu, S Yan, L Zhang, H Zhang and W Gao 2005 Efficient 3D Reconstruction for Face Recognition *Pattern Recognition* **38** 6 787–798
- [10] S Romdhani and T Vetter 2003 Efficient, Robust and Accurate Fitting of a 3D Morphable Model *Ninth IEEE International Conference on Computer Vision Proceedings IEEE* 59–66
- [11] C Tomasi and T Kanade 1992 Shape and Motion from Image Streams Under Orthography: A Factorization Method *International Journal of Computer Vision* **9** 2 137–154
- [12] J Fortuna and A M Martinez 2010 Rigid Structure from Motion from a Blind Source Separation Perspective *International Journal of Computer Vision* **88** 3 404–424
- [13] M Castel'an and E R Hancock 2006 A Simple Coupled Statistical Model for 3D Face Shape Recovery *IEEE 18th International Conference on Pattern Recognition ICPR 2006* **1** 231–234
- [14] M Song, D Tao, X Huang, C Chen and J Bu 2012 Three-Dimensional Face Reconstruction from a Single Image by a Coupled RBF Network *IEEE Transactions on Image Processing* **21** 5 2887–2897
- [15] K Punnam C and T Satya S 2015 3D Face Model Estimation based on Similarity Transform using Differential Evolution Optimization *Eleventh International Multi-Conference on Information Processing-2015 (IMCIP-2015)* Procedia Computer Science 54 621 – 630
- [16] D Fidaleo and G Medioni 2007 Model-assisted 3D face reconstruction from video *Lecture Notes in Computer Science* 4778 124–138
- [17] A K Roy C and R Chellappa 2003 Face reconstruction from monocular video using uncertainty analysis and a generic model *Computer Vision and Image Understanding* **91** 188–213
- [18] S Lee, K Park and J Kim 2011 A SfM-based 3D face reconstruction method robust to self-occlusion by using a shape conversion matrix *Pattern Recognit.* **44** 7 1470–1486
- [19] Boulbaba B A, Mohsen A, Liming C 2006 An Improved 3D Human Face Reconstruction Approach Based on Cubic Splines Models *Proceedings of the Third International Symposium on 3D Data Processing, Visualization, and Transmission (3DPVT'06)*
- [20] Seema C and Prajakta A S 2017 3D Face Reconstruction from 2D Images Using Data Mining

Algorithm *International Journal of Advance Research, Ideas, and Innovations in Technology*  
**3** 3 319-324

- [21] Christian B, Mauricio R, Mihai C, Radu O, Edilson de A, and Thiago O-S 2014 3D Face Reconstruction from Video Using 3D Morphable Model and Silhouette *Graphics, Patterns and Images (SIBGRAPI)*
- [22] Xiaoping H, Ying W, Feiyun Z and Chunhong P 2016 LEARNING-BASED FULLY 3D FACE RECONSTRUCTION FROM A SINGLE IMAGE *ICASSP Proceeding* 1651-1665
- [23] Shu Z, Hui Y, Ting W, Junyu D and Honghai L 2017 Dense 3D Facial Reconstruction from a Single Depth Image in Unconstrained Environment *Virtual reality* springer 1-10
- [24] Ira K-S and Ronen B 2011 3D Face Reconstruction from a Single Image Using a Single Reference Face Shape *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE* **33** 2 394-405
- [25] Ramasubramanian M and MA Dorai R 2012 Reconstruction of sensible 3D Face Counterpart From 2D Images-A hybrid approach *The International Journal of Research Journal of Computer Systems Engineering, ISSN 2250-3005*