Multimed Tools Appl (2018) 77:25253–25273
https://doi.org/10.1007/s11042-018-5783-1

A relative 3D scan and construction for face using meshing algorithm

Dayanand G. Savakar[1] **· Ravi Hosur**[2]

Received: 20 April 2017 / Revised: 15 December 2017 / Accepted: 9 February 2018 /
Published online: 3 March 2018

Abstract Three-dimensional object construction has seen a great deal of activity in the past decade, as has been pointed out in recent surveys, with the advancement of technology and easy availability of the depth sensors the 3D scanning technology has taken off. A wide range of commercial sensors such as Intel RealSense, Microsoft Kinect are being widely used for real time 3D capturing. Efficient 3D face scanning is one of the important areas of 3D scanning. 3D printer compatible texture supported scanning has a wide range of commercial applications. Such methods are also being used for 3D avatars and characterizations for games. Even though several commercial grade applications are available, most of the techniques suffer from background and light variations. Therefore an efficient face scanning technique is of extreme importance. In this paper we propose a 3D face scanning method based on Intel RealSense technology that combines 3D face detection, background segmentation and 3D mesh mapping to produce realistic 3D face model along with texture export. Further the models then can be manipulated using any 3D editing software along with texture and wireframe manipulation.

Keywords 3D image · RealSense · Depth · Segmentation · Texture · Projection · Meshing · Point cloud

✉ Dayanand G. Savakar
dgsavakar@gmail.com

Ravi Hosur
hosuravi@gmail.com

[1] Department of Computer Science, Dr P G Halakatti Post Graduate Centre, "Vachana Sangama", Rani Channamma University, Toravi, Vijayapur 586108 Karnataka, India

[2] Department of Computer Science & Engineering, BLDEA's V. P. Dr. P. G. H. College of Engineering & Technology, Vijayapur 586103 Karnataka, India

1 Introduction

Three-dimensional reconstruction of objects has become an important technique in the Industry, Archaeology analysis and Forensic analysis in Medical sciences with respect to the contexts of inspection for quality observations from different views or from unseen actions, reconstruction from raw data and for indulging knowledge, etc. Generally, a 3D model (Fig. 1) is essentially a set of vector vertices of connected triangles.

In the 3D model, the triangles create a mesh which is commonly referred to as a wireframe. A wireframe contains coordinates of the vertices and the joining line and does not account for any colors visuals where the color or the texture of each of the connected triangles is identified by a texture map.

In the proposed method, a 3D scan is essentially performed by defining a set of vertices on the 3D space and estimating the proximity of the points with respect to the points in the depth map. During the process, one of the biggest challenges is with a realistic scanning that accounts for the noisy depth data. Several filtering algorithms have been proposed in the past to filter depth data. However, disconnected point results in poor optimization of the data. To filter all these flaws, Intel RealSense technology offers an out of the box facial scanning process. In the work, we first understand existing techniques and the basic principles of face scanning technique offered by RealSense, the drawback and later elaborate on the proposed technique.

2 Related work

2.1 Literature survey

(Herrera et al. [9]) Proposes a method of 3D object reconstruction by combining dense SLAM, object detection and 3D segmentation by finding the dense scene structure, to provide shape information for pose estimation. The work reveals the advantage of depth sensors by combining images of joint type and the depth factor in order to have robust. The purpose of 3D tracking and reconstruction of objects incorporation of mirroring tech, aided by edge based segmentation provides much better results than color histogram.

(Zhongjun et al. [22]) Proposes a multi-depth GEM for modelling a 3D face by using the depth maps generically. The experiment carried gives better results compared to GEMs with a margin in recognition to the collected images.

(Zhou et al. [27]) Proposes the face recognition method in handling the variations which occur in pose of a face object during identification using Generic Elastic Models (GEMs) that extends to multi-depth by the help of different depth maps. By using multi-depths of a face so that the identification has given in better results with least correlation value. By using linear regression method each 1D face depth are rotated, rendered, and synthesized to 2D projection for their identification.

(Kaneko et al. [12]) Proposes a 3D edge detection technique depending on a decision tree constructed for a 3D depth image with better and faster results than machine learning.

(Xu et al. [24]) Proposes a method for constructing a full color 3D model by integrating the images captured by a 3D camera with an array of scenes differentiated by a small angle. The method also proposes a pre-processing filter where with a single step removes all dark borders that were captured along with the scene and also removes the noise from the scene without disturbing the quality vision of the scene of the captured.

Fig. 1 A 3D face model with texture

(Cheng et al. [3]) Proposes a method 3D constrained local model framework for face with the different landmark localization. With the experimental results, the method proves that the intensity of data and 3D feature fusion have improved the performance.

(Zhang et al. [26]) Proposes a least square depth estimation technique for the reconstruction of 3D from a captured 2D scene. In the process, number of textures is captured as the features of an image for estimating the depth on experimentation resulted in robustness property for the special depth cues for describing the relationship between 819 characteristics and depth of the scene.

(Dinc et al. [6]) Proposes a method for constructing a texture based 3D model for an image (3D) captured by MS Kinect camera system. The technique consider SIFT and SURF features, then constructs non-textured 3D model, mapped with color image an Delaunay triangulation. The experimental results show the construction process of 3D with no significant problems to existing criteria.

(Poon et al. [17]) proposes a method localizes the accuracy of the framework of a constructed 3D face by robusting the intensity of data and features of 3D depth histogram. The conduct of experiment has resulted in the significant variation in the performance with the localization of framework model involving defined constraints.

(Daribo et al. [5]) Proposes a technique that first represents ASMP by constructing non-overlapping sub-blocks by dividing the boundaries using arithmetic edge coding (AEC), and secondly to compress the dividing boundary. Experimental results shows that the boundaries identified of an object at both low and higher bitrates are preserved by enhancing the performance of depth based image processing applications.

(Dame et al. [4]) Proposes a least square estimation technique for reconstructing 3D image using 2D based on multi-textured features to filter texture gradients, variations and colors where the features are trained for estimating the depth. The method results are good to the scene depth information extraction and also advances in usage of less space of 2D images also bring the relationship between constraint characteristics and depth information.

(Ionescu et al. [11]) Proposes a new technique for gesture control with a Kinect camera. In the process IR based 3D is captured so that in turn helps in capturing 3D depth base images using space splicing method. In the process, the camera captures the gesture made from the user by detection, tracking and recognition of the depth image map. Finally they construct 1 3D ma of the object based on the sequence of images.

(Swash et al. [19]) Proposes a new method for constructing a 3D color model with depth control, robustness and real-time using Holoscopic 3D imaging camera. To avoid the disturbances in the recorded image, the technique built by pre-processing captured images is carried. Experimental result shows that the technique will help in removing the dark borders by computing efficient and fast real-time 3D color model.

(Manap et al. [15]) Proposes a new DILS algorithm in construction and representation of interview images based on disparity depth by selecting the areas of depth and its representation. The construction helps in reducing the complexity of 3D image and view point applications. Experimental results show that the method can be involved in synthesis of better quality inter-view images.

(Zujovic et al. [28]) Proposes a method for measuring the structural similarity measures between visual perceptions and textures obtained naturally. The measurement is completely based on local image statistics and point-by-point deviation between textures. These are executed with the extension of ideas of SSIM. The construction of the method is implemented using steerable filters decomposition and also by incorporating sub-band statistics. Experimental results prove that the method performs best for texture patch retrieval.

(Wang et al. [21]) builds a 3D morphable face model, which can be employed for image based facial performance capturing by obtaining low-frequency 3D information and rely solely on normal data from multiple views by using non-linear optimization technique for mapping 2D. The technique explains how to reduce the acquisition time to generate 3D construction and align semantic templates for matching.

(Park et al. [16]) Proposes a method which reveals the idea of 3D object recognition using the 3D edge detection method with the help of a decision tree generated for detecting the 3D edges where the tree is trained under supervised learning process for images and the relationships identified between pixels. Hence, the training the better results on conduct of the experiment for convex corner detection in 3D image faster than the conventional method.

(Xu et al. [23]) Proposes a method to solve the problems faced in conversion from 2D to 3D image and also to combine depth maps from graph-cuts sharp boundaries. The method is designed using fast watershed segmentation based on priority queue to generate constraints on depth map by replacing graph-cuts. Experimental results show that the process takes less time in producing good quality images.

(Wang et al. [20]) Proposes a method that can be used in semi-automatic 2D to 3D conversion by considering the depth values of neighbouring pixels. The process first identifies the pixels with the same depth values along with the nearer depth valued pixels for their depth estimation. Experimentally proved that the method's performance and speed are increased 5 times of an average compared to existing one.

(Ranipa et al. [18]) Proposes a method for estimating the depth as well as image restoration from the defocused results. The method first estimates depth using Subbarao method and then refine the depth image with a fast optimization and restoration technique using Weiner filter. Experimental results proved that the restored image obtained is still better under space-varying blur conditions.

(Kemelmacher-Shlizerman et al. [13]) method obtains an input a single image along with a single 3D reference model of a different person's face. Reconstruction methods from single images, i.e., shape-from-shading, require knowledge of the reflectance properties and lighting as well as depth values for boundary conditions, input image as a guide to "mould" a single reference model to reach a reconstruction of the sought 3D shape with accuracy and robustness are proved for the considered database images.

(Xiang-Cheng et al. [2]) Proposes a method of 3D reconstruction from uneven defocusing image to recover depth in the image. In the construction the sequence of images are generated using counter-fitter that are closely related with the radius of the blur circle. By calculating focusing and defocusing, and defining the parameters of focusing point establishes the relationship between defocus measurement and depth transformation of the image sequence. Experimental results proved that the method is valid for 3D reconstruction from uneven defocusing image.

(Lopez-de-Teruel et al. [14]) Proposes an algorithm for the reconstruction of indoor scene from captured image. The solution is constructed by the orientation of planes (horizontal and vertical) with the zone classification algorithm to divide them into patches of disjoint. With the combined techniques of feature extraction and augmentation of local colorization algorithms, reconstruction of a scene is achieved. Experimentally, depending on the extracted image segments and 3D reconstruction a single geometric projection is obtained.

(Zhang et al. [25]) Proposes a method for depth image based generation of virtual stereoscopic images by minimizing the distortion caused to symmetric smoothing. The method uses a technique of asymmetric smoothing to smoothen the sharp changes in depth at object boundaries, and hence reduces the geometric distortions. Experimental results prove the improvement in the quality of stereoscopic virtual images with depth quality.

(Fehn et al. [7]) Proposes a new approach of 3D TV in comparison to existing 2D digital TV framework. The work describes the application of DBIR technique, DBIR algorithm to generate high quality virtual stereoscopic views, and transmission of MPEG tools.

(Gadkari [8]) proposes an investigation method for a gray level co-occurrence matrix which is used in texture classification. In this method decision is made on the comparison between image textural features and image memory size for optimization where the image stored in different levels of compression is considered. Experimentally, proved that GLCM is a good discriminator.

From the above survey, it is found that numbers of techniques were proposed for the construction and reconstruction of 3D image from a 2D image of different collections. There were results the proved to be better in the process of constructing 3D from different depths, but the methods did not have sufficient proof of constructing a 3D image with texture transformation, correlation between color and texture. Most of the methods suffer from light variations occurred in the data at the background, that are taken as input in the construction of 3D. In the following sub-sections 2.2 and 2.3) we will discuss regarding scanning, mesh generations and 3D face model construction.

2.2 3D scanning process and mesh generation

With the idea of face scanning process, the complete 3D face scanning process of the 3D face is shown in Fig. 2. Firstly, depth map data captured by IR sensors is projected with RGB color camera data to obtain a background segmented perspective of the face. Once the scanning process starts, a set of 3D points is captured from the segmented frame. The coordinates of

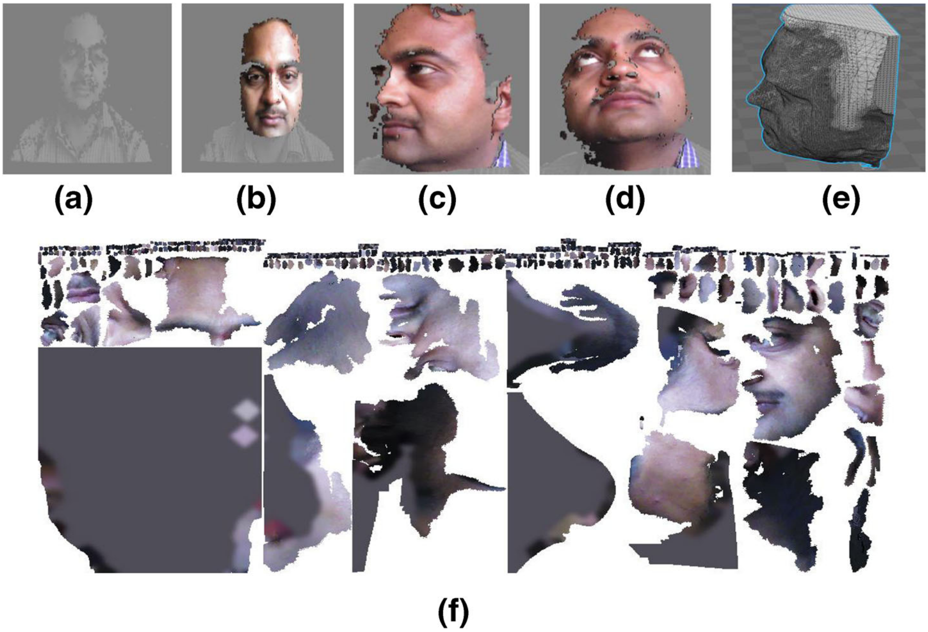


Fig. 2 Face 3D scanning process by RealSense for Fig. 1 (**a**) Depth map (**b**) Projection of color on the top of depth map (**c**) correlating facial points on depth data (**d**) Facial rotation for obtaining point cloud in 3D space (**e**) Point cloud triangulation mesh and (**f**) Texture export

these points are also known as point cloud. The points in the point clouds are connected through sets of lines forming triangle. Interconnection of all such points is also known as 3D mesh model.

The mesh model shown in Fig. 2e is zoomed to see the wire framed network of triangles and is presented in Fig. 3.

At, any given instance, the view presented to the user is 2D, however not all the facial points can be captured in a single view. Therefore once the scanning process starts, set of points are captured. These points serve as reference points. User needs to move his head in all directions (Fig. 2d) to keep accumulating the points from every part of the face. The new points are checked against the points being already acquired. As the rough point cloud and the mesh looks like Fig. 3, which does not elaborate the facial texture, it is important to extract the texture. Texture file is extracted as a 2D image file as shown in Fig. 2f. The image file essentially comprises of several isolated texture blocks. A metadata called mtl file is generated to map these textures with point cloud mesh which is stored in .obj file.

The texture is exported as a 2D image map, all the parts of the texture map can be retouched and different color or material can be applied on the model. Figure 4 shows various texture variation of the model. The overall process can be presented with the help of a simple block diagram as shown in Fig. 5.

2.3 3D spherical face model

Because of the interlinking of the point-cloud from user's facial pose variations is based on both depth and color data, the color data loses the actual intensity and color composition. The

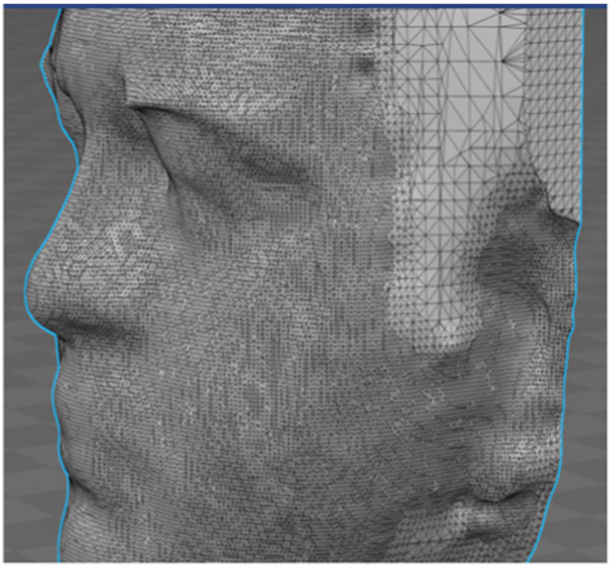


Fig. 3 Closer look at the cloud mesh of Fig. 2e

scanned model also suffers from loss of hair and head information as black hair tend to observe IR light and therefore gets missed in the scanning.

The other notable drawback with the scanning process is that, only half of the facial sphere surface is captured (essentially the frontal surface). A comprehensive view of the half sphere frontal face is as shown in Fig. 6.

Though such a model doesn't present any drawback for frontal facial manipulations, most of the applications like 3D printing, 3D avatars desire a complete head profile. Therefore the problem space then can be defined as to develop a Facial 3D scanning technique for the entire facial geometry with the ability to manipulate point cloud, mesh and texture. The generated model must also retain the realistic skin texture which is missing from the SDK's support. The method is elaborated in detail in the proposed system section.

3 Proposed system

From the above quantified details, the 3D image reconstruction is with the consideration of depth and texture factors. The Fig. 7 shows the two conventional strategies of constructing a

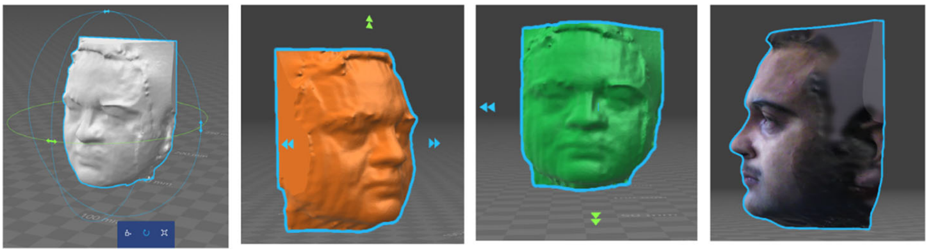


Fig. 4 3D face scan with texture retouch

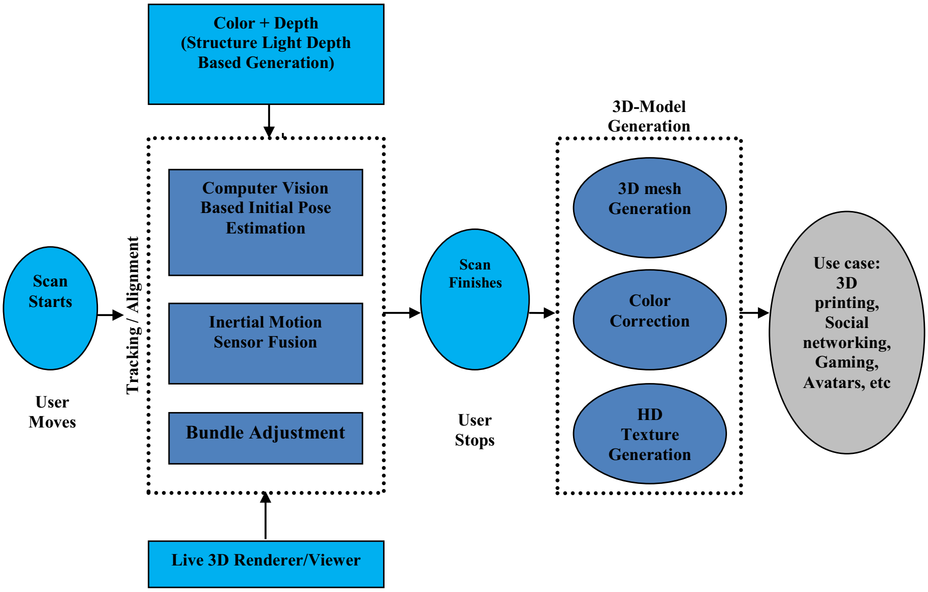


Fig. 5 General block diagram of 3D face scanning

scene object for scanning. Here, we will first discuss how the conventional 3D construction is performed and then elaborate our contribution to the architecture.

3.1 Conventional color map projection on depth map

In the first method, the RGB color stream is projected on the top of depth map imaging. Depth map imaging is a 2D image which essentially encodes the distance of all the points in an object from the observer. It is generated through the IR depth camera. The color map image is essentially the colour RGB 2D image. By thresholding depth map image and eliminating the distant objects, we can get a very accurate depth segmented 2D image. When this 2D image is combined with the color image, we obtain a segmented color image which only contains the foreground object.

As depth camera and color camera sensors are displaced by a small distance in the space, the points in both the data are displayed by small centre converging distance. RealSense provides ready interface for the projection and therefore the process of projection is omitted from the paper.

We can see from the result that such a projection results in far more noise in the facial region that the second strategy where color is projected over a world imaging which is generated out of projection from color over depth. However, in both these models, we can see immense amount of noise around the face region and occlusion of the hair region. We can also see that the facial part blob is much brighter in comparision to the rest of the body part. This is because a depth map adopts a distance based color scheme. Therefore if the scanning strategy asks the user to bring his head closer to the camera, rest of the body can be eliminated from the depth imaging by simple threshold based blob segmentation.

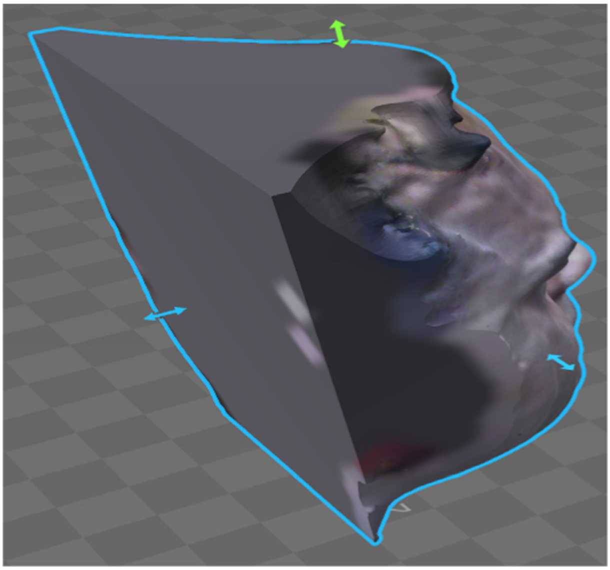


Fig. 6 3D face model with absent back sphere

3.2 Improving depth based segmentation for face using a facial mask

As we learnt that conventional technique is more suited for rigid objects and causes significant noise in the projected and segmented frame, we introduce facial mask. A mask can be generated and superimposed on the world image to consider only said part of the image for acquisition of the point cloud. Such a facial mask is shown on Fig. 8.

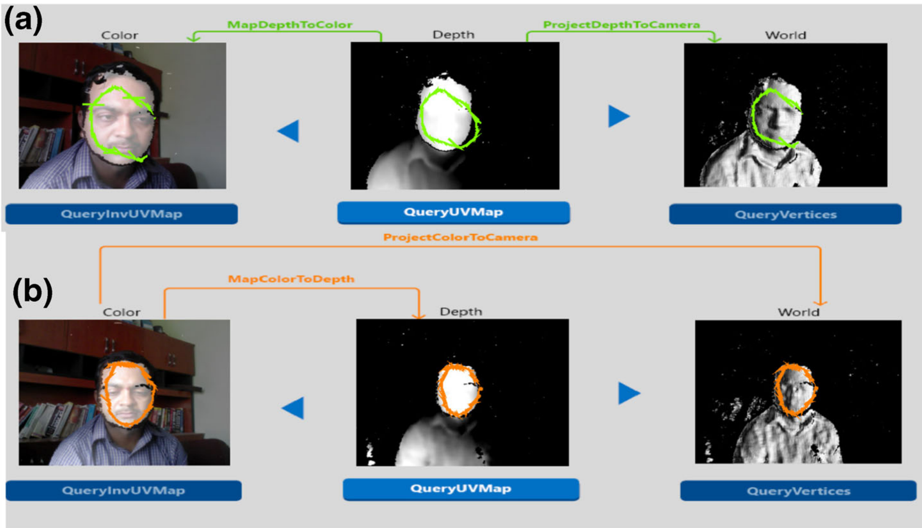


Fig. 7 Conventional projection technique (**a**) Color over depth projection (**b**) Color over (Color over depth) projection

Fig. 8 Facial mask on the top of color and depth imaging

We first use Voila Jone's face detection method on the 2D color image, an elliptical mask corresponding to the rectangular facial area in the RGB image is placed on the depth image and rest of the area around the face part is eliminated from the depth map. Then depth map is projected with the color map to obtain a noise free facial segmentation of the face region.

As absolute brightness of the depth image is related to near zero distance from the camera, the resultant mask brightness is close to unity but not unity. Observation shows that average normalized mask intensity is about 0.88. When this mask is multiplied with the segmented facial model, the color brightness of the final texture tends to degrade. This is also evident from the final scanned rendering of the face shown in Fig. 1. However if we first threshold the depth map and then project the depthmap over the color image, the resultant world image is literally a background separated image without noisy point cloud around the face. Background separated world image is shown in Fig. 9.

It is clear that background segmentation yields much better world image for processing than non background segmented image. Therefore we append this block to Fig. 5 to obtain a much better world image for processing. The modified block diagram of producing background separated and noise-free world image is presented in Figs. 10 and 11.

We can clearly observe the improvement in overall head profile of present and proposed system. In the above block diagram, new blocks that are being added with the present system are shown as 3D rectangles. Having being able to eliminate the point cloud noise around the face, we now need to focus on the other part of the problem, i.e. to extract the entire facial model as a complete sphere, followed by dividing the sphere with connected mesh and overlapping the texture on the top of the mesh. In a 3D rendering context, points in different face angle views must be connected as a single connected object to offer smooth rendering. This is achieved in any 3D object by typically creating a polygonal mesh S = {V, E} which is

Fig. 9 Background segmented world image (**a**) Depth image (**b**) Color Image (**c**) Background separated image

Fig. 10 Non background separated world image

defined by a set V of vertices Vi and a set of edges E as shown in Fig. 2 of [21]. In the next section we will discuss about the mesh generation from our segmented world image.

3.3 3D mesh generation

Even though, there are several previous works on 3D mesh generation for 3D object scanning, method proposed by [21] is most suitable one in terms cleaner mesh generation. Instead of aligning and trying to fit a polygonal mesh on the entire face or 3D model, instead certain facial features are aligned. This technique introduced as feature based registration which is applicable for matching facial point on the face. However, the effectiveness of the model

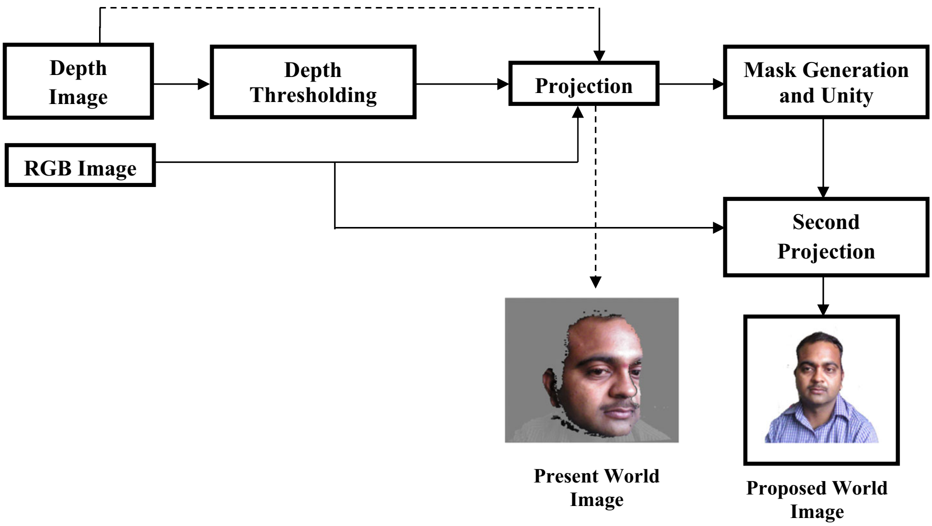


Fig. 11 Present v/s proposed pre-processing and world image generation

depends on the feature alignment. Fortunately, all the faces are of similar features. There are key points in the face which varies from face to face.

If we consider a generic face model and superimpose key facial features from the scanned 3D model on the top of it, then we may get a more complete facial model. Figure 12 shows the strategy of transferring key facial points from scanned 3D face to generic 3D face.

The key facial feature points are also known as landmark points. Therefore the next strategy is to obtain facial landmarks from the scanned face and blend it on the generic face. However, the problem with such a technique is that everyone has a different shape of the face defined by the facial boundary geometry. Different mesh geometry is not considered in [21] and therefore the base model is fixed. But we first reform the base model based on facial landmark features. This includes roundness of the face, width of the forehead, length of the jaw, and so on. Hence only transfer of facial landmarks will result in a deformed scan as the resultant geometry will differ from the actual geometry. We can see from the above image that such a transform often results in faces that are way too different from the scanned face. Therefore alternative strategies are required for the transform.

In order to overcome the drawback of Fig. 13, we perform facial boundary tracking. Facial boundary tracking is performed by considering the boundary points of the generic 3D face as model parameters and matching the points of the scanned face with the same. Once the facial boundary is extracted, the generic 3D face boundary points are adjusted accordingly. We adopt the landmark tracking and boundary tracing in the depth space as the model parameters are entirely defined by the positional vectors and not the color space. As the images are already segmented and world imaging doesn't have any other parts other than the faces, the depth data based tracking results in faster and more accurate tracing (Figs. 14 and 15).

Once the facial landmark points and the boundary are extracted, we create a 3D mesh with these points. Instead of merely transforming the facial landmarks on the top of the generic face, which is being used by some of the current techniques, we first apply facial geometry transform and then follow this by transforming facial texture around the geometry over to the generic face rather than simply transferring the landmark area. Proposed technique is depicted in Fig. 16.

Overall process is explained with a help of Fig. 17, which demonstrates the difference between the present and the proposed system. The updated and introduced blocks are all depicted by 3D rectangular block. For optimizing the mapping process, user is required to move his face slowly in UP, DOWN, RIGHT and LEFT directions, where the model and actual feature points are compared and mapped.

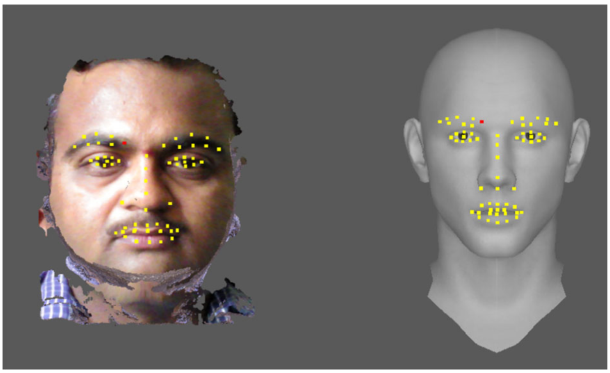


Fig. 12 Facial landmark points

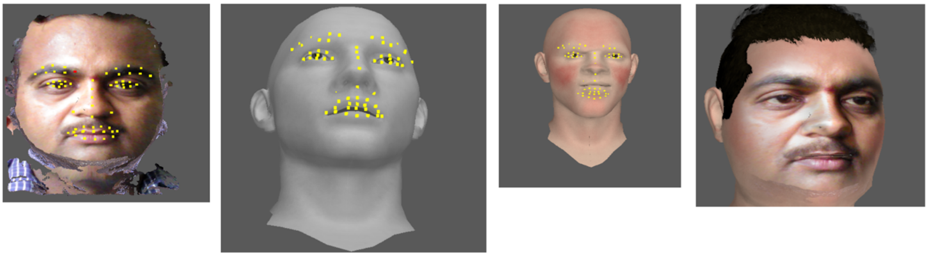


Fig. 13 Single step facial landmark transform

3.4 Texture mapping

Once facial geometry is fitted on the mesh geometry, the extracted texture map must be applied on the geometric mesh model. A Cylindrical transform for texture mapping is suggested in [10]. However, with this method, the texture appears unrealistic and needs to be smooth, because 3D depth map contains larger volume by definition than 2D texture map. Therefore we divide the texture mapping in two distinct parts explained below. We not only adopt a facial geometry translation, but at the same time we divide the texture mapping to depth and color mapping. The advantage of such mapping is such that the user can retouch the generated model's mesh points along with the texture points once the final scanned face is generated. The texture transfer is shown in Fig. 18.

The process separates the segmented RGB map, depth map and initially starts by mapping the depth map data on the top of generic face model. It is then followed with feature mapping, color transform and skin map model respectively. At the end of the transform, we obtain a face where color blend difference is visible. In order to obtain a much realistic face model similar to the skin texture of the actual face, we perform lighting readjustment in order to obtain final 3D face model.

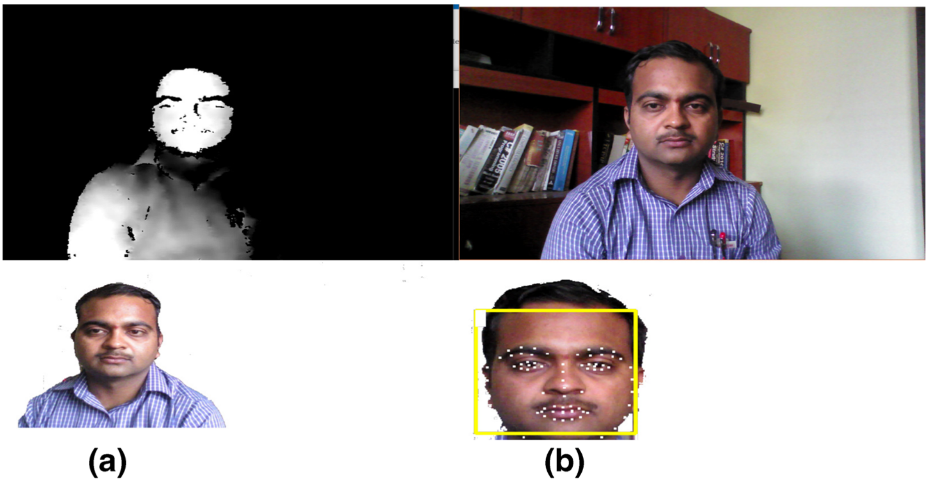


Fig. 14 Facial landmark with face boundary (**a**) Segmented World Image (**b**) Face landmark with boundary

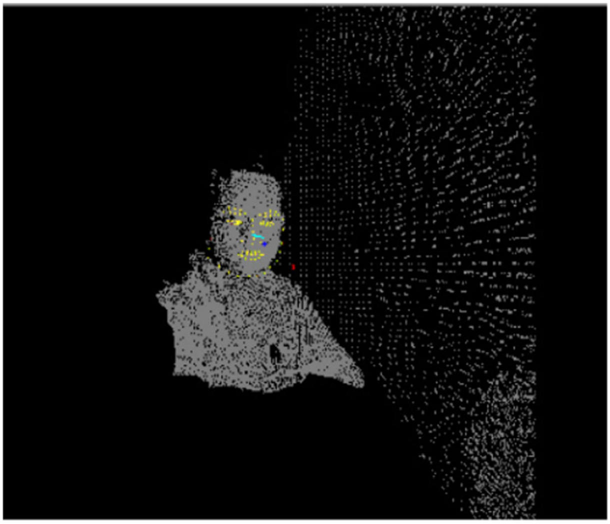


Fig. 15 Facial landmark tracking in the depth space

4 Results and discussion

From the above stated methodology of constructing a 3D face, the following models are constructed in Figs. 19 and 20.

Shape correlation [7] is defined as the percentage of similarity between the edges of a frontal face image with that of edges extracted from the frontal view of the 3D image. The edge detection helps in identification and construction of the 3D face with proper boundaries in

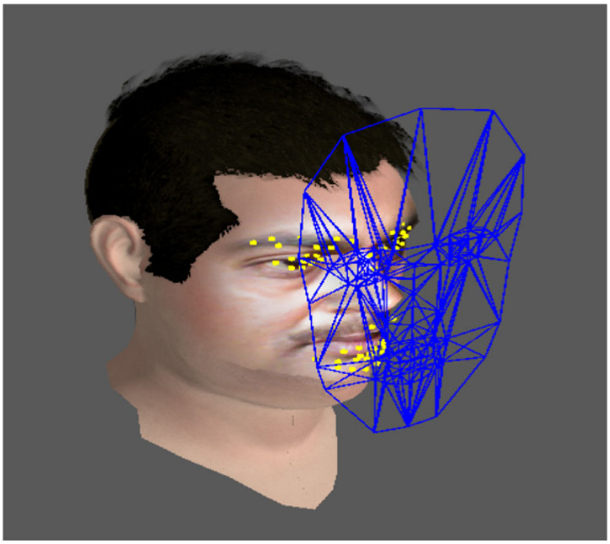


Fig. 16 Facial geometry transform

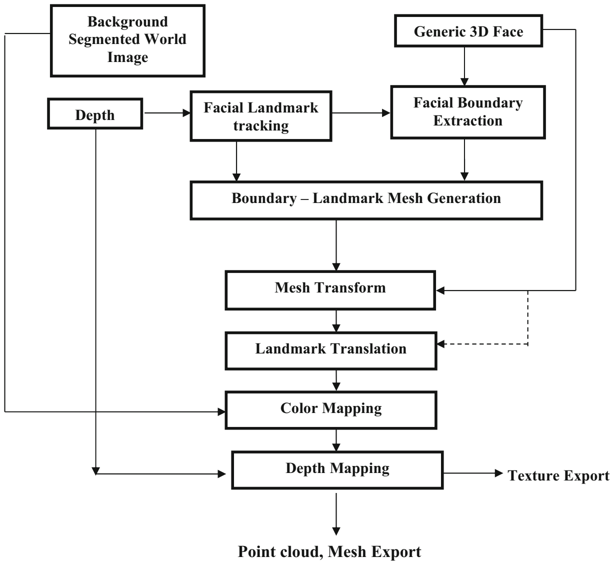


Fig. 17 Proposed Facial geometry projection block diagram

circumstances like background of the capturing image, and the areas of the face to be filtered from a capturing image containing elements like hair. The results are tabulated in the Graph 1 in comparison to the techniques of 2D interpolation method and absolute method used for construction. Shape correlation also represents the homogeneity in the triangulation and output mesh model. Deviated edges represent unrealistic modelling. We can clearly see that the performance of proposed system is better than the existing systems in most of the test cases barring user no. 4 which is because of the significant facial hair that the user had (Fig. 20b). The compared correlation factor performance clearly demonstrates the improvement achieved by proposed system over popular current state of art.

Texture correlation on the same way is defined as the similarity between the GLCM texture values [8] of a static face image with 2D image extracted from frontal view of the 3D model.

Fig. 18 Texture Transfer from segmented world image to generic face model

Fig. 19 Face scanning of an Indian male

Similarly in the context of constructing a 3D face, when considered with the texture the face is correlated with the eyes, nose, hairs on the face and head. The texture features can also be represented by structural textures as suggested by [25]. These textures are constructed with better results in comparison with the 2D and absolute 3D construction methods where the proposed technique removes the face is constructed without hairs in the head, also the images captured with or without hairs on the face. The results are tabulated in Graph 2.

Even though, our work focuses on the capturing real facial 3D data and modelling it into a synthetic 3D face, it is essential to analyze the performance of the method with standard datasets for better understanding of the performance.

We used [1] for evaluation of the registration error of the proposed method and compared the same with 2D interpolation and Absolute 3D scan methods. Absolute 3D scan is the ground data provided with the dataset. We used 2D interpolation with Median filter to fill the gaps after feature mapping. In the proposed system, as the database images contains only one view of the face, the view was rotated around center point and the landmark points are adjusted. Sample output of the three is shown below (Fig. 21).

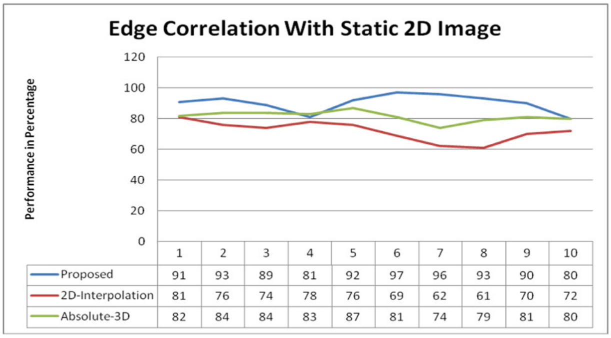
The performance is shown in the graph below:

In the process, we compared the distances between the positions of the landmark in the resulting textured transformed image with that of the ground truth data and calculated the average normalized distance. Registration accuracy is given by 1-Registration error. From the Graph 3, we observe high error in 2D method due to dilution of facial curve and applying median filter. The proposed method on the other hand relies on accurate landmark tracking and texture mapping between polygonal mesh and facial landmark points. Error was minimum for pure frontal faces and maximum for side faces. Hence, our method can be relied for the faces that can be captured from the front end to obtain higher accuracy.

The results obtained for reconstrucing 3D face with 0.81 accuracy using SSIM [9] method has overcome with a small amount by our technique with an averge result of 0.872. The 3D reconstruction [27] based on light intensity our technique works better in all variations of light.

(a) 3D Scanning of an Indian female **(b)** 3D scan of an Indian teenager with facial hair

Fig. 20 **a** 3D Scanning of an Indian female (**b**) 3D scan of an Indian teenager with facial hair

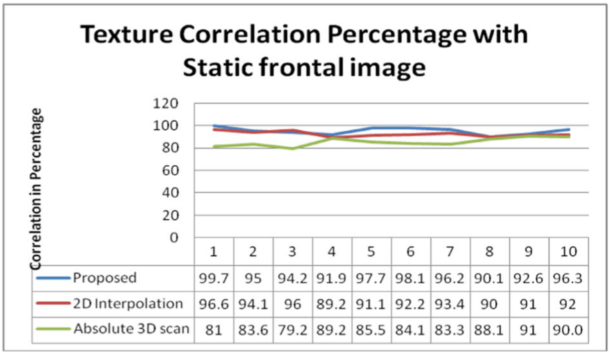


Graph 1 Result of comparison with texture performance correlation performance between proposed system and the present state of art

Even for localizing the landmark points [3] our method generated a mesh framework to identify the landmark points with more or uneven light intensity with varying depth so that the performance can be improved with time and accuracy. 3D texture registration [6] were achieved with initial non-textured 3D model, whereas in our method both texture as well as non-textured can be registered with different depths captured from camera. Also the method [13] suggests that the construction of 3D face is obtained by using a single refence image where our method reconstructs a 3D face by capturing it in realtime from various views and as well as from stored dataset [1] which proves as an improvement to the work carried. Hence, our method is effective and flexible for both stored as well as realtime scanning methods of a face to reconstruct its 3D model view to the existing 3D reconstruction methods.

5 Conclusion

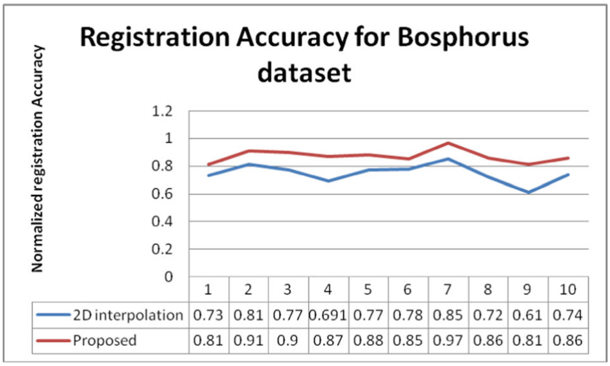
3D scanning is one of the major areas of innovation in the decade. Even though several techniques were proposed earlier for 3D scanning, introduction of low cost fusion imaging has accelerated the development in 3D scanning area. This has lead to the research as well as commercial solution in 3D printing and scanning. Even in the presence of various APIs, tools, and low cost devices the efficient extraction of object 3D model under real-time scanning condition still remain a big challenge. After evaluating several past techniques we concluded that there is an immense



Graph 2 Result of comparison with edge correlation

Fig. 21 Comparison of 2D facial registration with median filter and proposed technique for [1]. **a** Actual texture mapped image (**b**) 2D interpolation with median filter **c**: Proposed

possibility of improvement in the 3D scanning field. The facial scan presents greater challenge due to several traits like facial hair, hair, and so on which are traditionally late respondent to fusion imaging. A combination of restoration, synthesis and interpolation is proposed to produce more realistic scan. Our technique provides an average improvement over current state of art that can be seen in results obtained (shown in Graphs 1 and 2). We also evaluated our feature mapping based frontal facial 3D segmentation against standard Bosphorus dataset and compared the technique with popular texture mapping and 2D interpolation techniques. Result shows that our technique of morphing facial features into a standard model yields lesser holes after texture transfer and



Graph 3 Normalized registration accuracy

therefore is much smoother. Also the technique allows significant post processing capabilities where facial expression can be synthesized after face mapping. This makes our technique not only a good choice for 3D facial scanning but also facial model post processing.

References

1. Alyüz SN, Dibeklioğlu H, Çeliktutan O, Gökberk B, Sankur B, Akarun L (2008) Bosphorus database for 3D face analysis. The First COST 2101 Workshop on Biometrics and Identity Management (BIOID 2008), Roskilde University, Denmark, 7–9 May 2008
2. Chen X-C, Sheng Y, Wang Y-J (2007) Research on 3D shape reconstruction using uneven defocusing model. IEEE International Conference on Mechatronics and Automation (ICMA), pp 2326–2331
3. Cheng S, Zafeiriou S, Asthana A, Pantic M (2014) 3D facial geometric features for constrained local model. IEEE International Conference on Image Processing (ICIP), pp 1425–1429, October 2014
4. Dame A, Prisacariu VA, Ren CY, Reid I (2013) Dense reconstruction using 3D object shape priors. In: The proceedings of the IEEE conference on computer vision and pattern recognition, pp 1288–1295
5. Daribo I, Florencio D, Cheung G (2014) Arbitrarily shaped motion prediction for depth video compression using arithmetic edge coding. IEEE Trans Image Process 23(11):4696–4708
6. Dinc S, Sigdel M, Dinc I, Sigdel MS, Fahimi F, Aygun RS (2014) Depth-color image registration for 3D surface texture construction using kinect camera system. In: IEEE SOUTHEASTCON, pp 1–6
7. Fehn C (2004) Depth-image-based rendering (DIBR), compression, and transmission for a new approach on 3D-TV. International Society for Optics and Photonics Electronic Imaging, pp 93–104
8. Gadkari D (2004) Image quality analysis using GLCM. Electronic Theses and Dissertations, Paper 187
9. Herrera JL, del-Blanco, CR, Garcıa N (2015) Edge-based depth gradient refinement for 2D to 3D learned prior conversion. An IEEE 3DTV-conference: the true vision-capture, transmission and display of 3D video (3DTV-CON), pp 1–4
10. Hwang J et al (2012) 3D face modeling using the multi-deformable method. Sensors 12(10):12870–12889
11. Ionescu D, Suse V, Gadea C, Solomon B, Ionescu B, Islam S (2013) A new infrared 3d camera for gesture control. IEEE International conference on Instrumentation and Measurement Technology Conference (I2MTC), pp 629–634
12. Kaneko M, Hasegawa T, Yamauchi Y, Yamashita T, Fujiyoshi H, Murase H (2015) Fast 3D edge detection by using decision tree from depth image. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp 1314–1319
13. Kemelmacher-Shlizerman I, Basri R (2011) 3D face reconstruction from a single image using a single reference face shape. IEEE Trans Pattern Anal Mach Intell 33(2):394–405
14. Lopez-de-Teruel PE, Ruiz A, Fernandez L (2006) Efficient monocular 3D reconstruction from segments for visual navigation in structured environments. IEEE 18th International Conference on Pattern Recognition (ICPR 2006), pp 143–146
15. Manap NA, Soraghan JJ, Petropoulakis L (2013) Depth image layers separation (DILS) algorithm of image view synthesis based on stereo vision. IEEE International Conference on Signal and Image Processing Applications (ICSIPA), pp 61–66

16. Park A, Kim J (2012) GPU accelerated view synthesis from multiple RGB-D images. 19th IEEE International Conference on Image Processing (ICIP), pp 573–576
17. Poon G, Yeung Y-Y, Pang W-M (2014) Enabling 3D online shopping with affordable depth scanned models. IEEE International Conference on Smart Computing (SMARTCOMP), pp 150–155
18. Ranipa KR, Joshi MV (2011) A practical approach for depth estimation and image restoration using defocus cue. IEEE International Workshop on Machine Learning for Signal Processing (MLSP), pp 1–6
19. Swash MR, Aggoun A, Abdulfatah O, Li B, Fernandez JC, Alazawi E, Tsekleves E (2013) Pre-processing of holoscopic 3D image for autostereoscopic 3D displays. IEEE International Conference on 3D Imaging (IC3D), pp 1–5
20. Wang D, Liu J, Ren Y, Ge C, Liu W, Li Y (2012) Depth propagation based on depth consistency. IEEE International Conference on Wireless Communications & Signal Processing (WCSP), pp 1–6
21. Wang Z, Grochulla M, Thormahlen T, Seidel H-P (2013) 3D face template registration using normal maps. 2013 International Conference on 3DTV-Conference, IEEE
22. Wu Z, Li J, Hu J, Deng W (2015) Pose-invariant face recognition using 3d multi-depth generic elastic models. 11th IEEE international conference and workshops on automatic Face and Gesture Recognition (FG), vol. 1, pp 1–6
23. Xu X, Po L-M, Cheung K-W, Ng K-H, Wong K-M, Ting C-W (2012) Watershed and random walks based depth estimation for semi-automatic 2D to 3D image conversion. IEEE International Conference on Signal Processing, Communication and Computing (ICSPCC), pp 84–87
24. Xu X, Po L-M, Cheung C-H, Feng L, Cheung K-W, Ting C-W, Ng K-H (2014) Adaptive block truncation filter for MVC depth image enhancement. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp 544–548
25. Zhang L, Tam WJ (2005) Stereoscopic image generation based on depth images for 3D TV. IEEE Trans Broadcast 51(2):191–199
26. Zhang L, Chen T, Sun H, Zhao Z, Jin X, Huang J (2014) The image depth estimation based on multi-scale texture features and least-square method. 12th IEEE International Conference on Signal Processing (ICSP), pp. 816–820
27. Zhou Y, Guo H, Fu R, Liang G, Wang C, Wu X (2015) 3D reconstruction based on light field information. IEEE International Conference on Information and Automation, pp 976–981
28. Zujovic J, Pappas TN, Neuhoff DL (2013) Structural texture similarity metrics for image analysis and retrieval. IEEE Transactions on Image Processing, vol. 22, no. 7, July 2013

Dr. Dayanand G. Savakar (Date of birth 02-10-1968), having 26 years of teaching and 12 years of research experience, he has completed his B.E degree from Karnataka University Dharwad, Post Graduation from Birla Institute of Technology and Ph.D from Visvesvaraya Technological University., Belgaum, India. He is now working as Professor, in the Department of Computer Science, Post Graduate Centre, Rani Channamma University, Vijayapur-Karnataka, India. He has published more than 55 research articles in international journals/conferences. Currently he is guiding eight Ph.D candidates. He served/serving as a member of various boards of several universities. His areas of interests are – Image Processing, Pattern recognition and Information security.

Mr. Ravi Hosur (Date of birth 22-08-1983), having 10 years of teaching and 3 years of research experience, he has completed his B.E degree from Visvesvaraya Technological University Belagavi, Post Graduation (M.Tech) from Visvesvaraya Technological University Belagavi and pursuing Ph.D from Visvesvaraya Technological University., Belagavi, Karnataka (India). He is now working as Assistant Professor, in the Department of Computer Science and Engineering, BLDEA's Vachana Pitamaha Dr. P.G. Halakatti College of Engg. & Tech., Vijayapur-Karnataka, India. He has published more than 3 research articles in international journals/conferences. His areas of interests are – Image Processing, Pattern recognition and Information security.