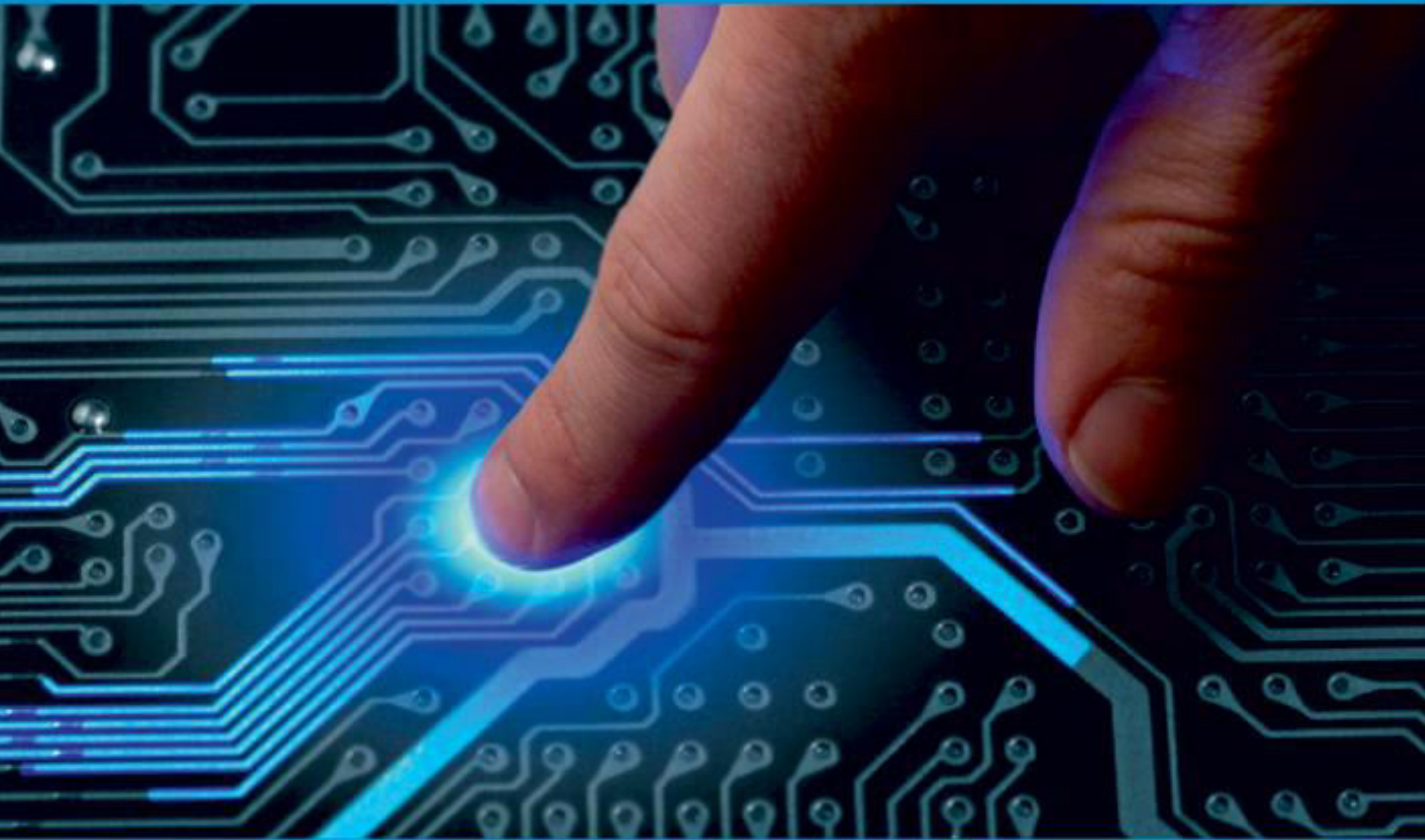
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 10, Issue 6, June 2022

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.165

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006082 |

Avatar Creation Using Face Reconstruction and Rendering

G.Iswarya [1], R.Sharmila[2], K.Shuba Shwetha[3], M.Varshini[4], M .Varthihaharsni[5]

Assistant Professor, Dept. of C.S.E, Saranathan College of Engineering, Trichy, India[1]

U.G Students, Dept. of C.S.E, Saranathan College of Engineering, Trichy, India [2,3,4,5]

ABSTRACT: A simple method that reconstructs the 2D face in a 3D plane while also providing dense alignment is proposed. A UV position map that stores the 3D shape of a face in a UV plane is created, it is regressed from a single 2D image using a simple Convolutional Neural Network (CNN). A weight mask is incorporated into the loss function, which is trained to improve the network's performance. This method does not require any prior face models and can reconstruct both the entire facial geometry and semantic meaning of the 2D face. Although approaches based on the 3D Morphable Face Model (3DMM) can reconstruct accurate 3D faces from single photos, the topology of the 3DMM mesh differs from that utilised in most games which requires a large amount of face texture data for training, and creating such datasets is time-consuming and difficult. Furthermore, a dataset acquired in a lab setting may not be able to properly adapt in the real world. To address this, a system that involves the following is proposed: A low-cost technique that acquires facial textures, A shape transfer algorithm that converts the shape of a 3DMM Face Mesh to Game Mesh, and A new pipeline that trains 3D game face reconstruction networks. The suggested method can produce detailed, vibrant game characters which are similar to the input portrait. It can also remove the effects of lighting and occlusions. Experiments on several datasets reveal that the system beats previous state-of-the-art methods for both reconstruction and alignment.

KEYWORDS: 3D Face Reconstruction, Rendering, Mesh, Texture Map

I. INTRODUCTION

An avatar is a graphical representation of a user, user's character, persona. Avatars can be two-dimensional icons in internet forums and other online communities, where they are also known as profile pictures, userpics or picons. Alternatively, an avatar can take the form of a three- dimensional model, as used in online worlds and video games. Avatar creation is an icon or figure representing a particular person like in a video game or in a cartoon movie. Perhaps an avatar can be created for the user when playing games. Avatars on Internet forums serve the purpose of representing users and their actions, personalising their contributions to the forum, and may represent different parts of their persona, beliefs, interests or social status in the forum.

A. Deep Learning

Deep learning is a subset of machine learning, which essentially deals with a neural network with three or more layers. These neural networks attempt to simulate the behaviour of the human brain-albeit far from matching its ability-allowing it to learn from large amounts of data. While a neural network with a single layer can still make approximate predictions, additional hidden layers can help to optimise and refine accuracy.Deep neural networks consist of multiple layers of interconnected nodes, each building upon the previous layer to refine and optimise the prediction or categorization. This progression of computations through the network is called Forward Propagation. The input and output layers of a deep neural network are called Visible Layers. The input layer is where the deep learning model ingests the data for processing, and the output layer is where the final prediction or classification is made. Deep learning utilises both structured and unstructured data for training. Practical examples of deep learning are Virtual assistants, vision for driverless cars, money laundering, face recognition and many more.

B. Position Map Regression Network

Position Map Regression Network (PRN) is self-supervised, jointly learning an appropriate geometric representation, a key point detector that finds points in common between partial views, and key point-to-key point correspondences. PRN predicts key points consistently across views and objects. Furthermore, the learned representation is transferable to classification.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006082 |

C. 3D Face Reconstruction

3D face reconstruction is the task of reconstructing a face from an image into a 3D form (or mesh). The 3D face reconstruction methods, to recover the 3D facial geometry under unconstrained situations from 2D images, are roughly classified into two categories: 3D Morphable Model (3DMM) fitting based method and End-to-End deep convolutional neural network (CNN) based method.

1) 3D Morphable Model

3D Morphable Model (3DMM) is a statistical model of 3D facial shape and texture in a space where there are explicit correspondences. The morphable model framework provides two key benefits: a point-to-point correspondence between the reconstruction and all other models, enabling morphing, and modelling underlying transformations between types of faces (male to female, neutral to smile, etc.,).

2) Convolutional Neural Network

A Convolutional Neural Network, also known as CNN or ConvNet, is a class of neural networks that specialises in processing data that has a grid-like topology, such as an image. A digital image is a binary representation of visual data. It contains a series of pixels arranged in a grid-like fashion that contains pixel values to denote how bright and what colour each pixel should be. The human brain processes a huge amount of information within the second an image is seen. Each neuron works in its receptive field and is connected to other neurons in a way that covers the entire visual field. Just as each neuron responds to stimuli only in the restricted region of the visual field called the receptive field in the biological vision system, each neuron in a CNN processes data only in its receptive field as well. The layers are arranged in such a way that they detect simpler patterns (lines, curves, etc.,) and more complex patterns (faces, objects, etc.,) as well. By using a CNN, one can enable sight to computers.

3) Convolutional Layer

The convolution layer is the core building block of the CNN. It carries the main portion of the network’s computational load. In this model, the convolutional layer is used as an encoder. This layer performs a dot product between two matrices, where one matrix is the set of learnable parameters otherwise known as a kernel, and the other matrix is the restricted portion of the receptive field. The kernel is spatially smaller than an image but is more in-depth. This means that, if the image is composed of three (RGB) channels, the kernel height and width will be spatially small, but the depth extends up to all three channels.

4) Transpose Layer

The transposed layer is also known as the Deconvolutional layer. A deconvolutional layer reverses the operation of a standard convolutional layeri.e. if the output generated through a standard convolutional layer is deconvolved, the original input is obtained. The transposed convolutional layer is similar to the deconvolutional layer in the sense that the spatial dimensions generated by both are the same. Transposed convolution does not reverse the standard convolution by values, but rather by dimensions only.

D. UV Space

UV space or UV coordinates which is a 2D image plane parameterized from the 3D surface, has been utilised as a way to express information including the texture of faces (texture map), 2.5D geometry (heightmap), 3D(geometry- image), and the correspondences between 3D facial meshes. UV space to store the 3D position of points from the 3D face model aligned with the corresponding 2D facial image.

E. UV Position Map

UV position map is a 2D image recording the 3D coordinates of a complete facial point cloud and at the same time it keeps the semantic meaning at each UV place. UV mapping is the 3D modelling process of projecting a 2D image to a 3D model's surface for texture mapping.

F. Differentiable Rendering

Differentiable Rendering techniques have been introduced to face reconstruction recently. It bridges the gap between 2D and 3D by allowing 2D image pixels to be related to the 3D properties of a scene. A differentiable renderer (DR) explicitly models the relationship between changes in model parameters and image observations. Differentiable rendering is a novel field which allows the gradients of 3D objects to be calculated and propagated through images.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006082 |

G. Shape Transfer

Shape transfer aims to transfer shapes between two meshes. To generate a full head model instead of a front face only, shape transfer is used to transfer a 3DMM mesh to a head mesh with game topology. Non-rigid Iterative Closest Point (Non-rigid ICP) algorithm is the typical method for a shape transfer task, which performs iterative non-rigid registration between the surfaces.

H. Mesh

Meshes are a feature of 3D graphics. They can be used in video games and in other software that models a three-dimensional environment. They do not have anything to do with game design except that they are building blocks for making visual elements of a game.

I. Loss Function

Loss function is used to minimise the distance between rendered face image and input face photo, and the distance between refined texture map and ground truth texture map.

II. LITERATURE SURVEY

Anh Tuan Tran, Tal Hassner, Iacopo Masi, Eran Paz, Yuval Nirkin, Gerard Medioni [1] discussed about the Existing single view face reconstruction has a major disadvantage of inconsistency in face of occlusions or non-frontal view inputs. A system with the capability of 3D face reconstruction under extremely adverse conditions and occlusions is built here, as the serendipity of the pictures that are fed might not always be favourable. A layered approach which decouples estimation of lower levels from its intricate details, inspired by the concept of bump mapping and it is used to represent complete face representation and their vertices allow for easy indexing of symmetric face regions. A deep encoder-decoder CNN is used to estimate the depth of the bump maps. Hence, the CNN, becomes capable of predicting the occluded regions of the face as well. It is shown that this method provides detailed 3D rendering of faces in an accurate manner compared to other conventional 3D face reconstruction methods in a single view. Aaron S. Jackson, Adrian Bulat, Vasileios Argyriou, Georgios Tzimiropoulos [2] explained 3D-face reconstruction and dense face alignment is an indispensable construct in Computer Vision, which is dealt with here. The face reconstruction from the 2D images of human faces that don't have a favourable serendipity, often requires methods that work with intricate details and dense correspondence of the pictures that are fed to the network. Such methods result in using complex and inefficient pipelines for model building and fitting. On account of minimising the inefficiency and complexity, a simple CNN is trained with suitable 2D image dataset and 3D facial models and scans. Anil Bas, Patrik Huber, William A.P. Smith, Muhammad Awais, Josef Kittler [3] determined the 3DMM is a conventional and very first model developed for face reconstruction and recognition. This model is trained with disparate datasets of faces along with the UV coordinates that are required to plot them in a 3D plane. The UV coordinates that are fitted with the 3DMM are known as the 3DMM parameters. The transformation of the input 2D image into a 3D module with the help of 3DMM is discussed here. The model is fitted with the CNN, which is an extension of the original spatial transformer network where interpretation and normalisation of 3D pose changes and self-occlusions are possible. The proposed 3DMM-STN has two parts: Localiser network and Grid generator network. The beneficial part of the model is the localisation part, that is trained to fit the 3DMM to a given 2D single input image. Chao Ye, Nanfeng Xiao[4] examined that 3D-face reconstruction and dense face alignment is an indispensable construct in Computer Vision. Creating position maps is a convenient method to reconstruct faces and align them. The novel upsampling block known as the shortcut upsampling block is being used for creating position maps. A CNN with residual and upsampling blocks is built which is referred to as the PRN. One of the important aspects of using the PRN is that it uses the position maps in reference from the 300W-LP dataset (dataset that is used for training) to represent 3D faces. Hence, a PRN consisting of few parameters and a better generality trait is created. At first, the residual blocks of the PRN are used to extract and regress the feature maps from the 2D face image preceded by the construction of position maps by the proposed shortcut-upsampling blocks of the PRN. Yao Feng, Fan Wu, Xiaohu Shao, Yanfeng Wang, Xi Zhou [5] The indispensable constructs in Computer Vision are 3D-face reconstruction and dense face alignment, which are dealt with here as well. Creating position maps is a convenient method to reconstruct faces and align them. A CNN wielding encoder and decoder blocks is built which is referred to as the PRN. One of the important aspects of using the PRN is that it uses the position maps that are extended from the 300W-LP dataset (dataset that is used for training) to represent 3D faces. Hence, a PRN with fewer parameters and better generality is created. Jiangke Lin, Yi Yuan, Zhengxia Zou[6] The conversion of the reconstructed face with dense face alignment into a game character (AVATAR), is dealt with here. Deep3DFaceReconstruction repository, that has the Facebook constructed CNN, to reconstruct faces along with dense alignment of it, is used. A shape reconstructor (CNN), trained with the BFM.mat files from the Deep3DFaceReconstruction repository, is used to predict the 3DMM and pose coefficients and reconstruct faces with dense alignment. Preceding which, a shape transfer module, converts the reconstructed face mesh to game mesh while preserving the base reconstructed face topology. Tianyang Shi, Yi Yuan, Changjie Fan, Zhengxia Zou, Zhenwei Shi, Yong

Liu[7] Automatic creation of in-game characters by replicating the game environment is dealt with here. A bone-driven model that predicts a set of facial parameters with a clear physical significance thereby producing a 3D profile for the given input, is built. In addition to controlling individual face attributes with each face component like position, orientation and scale; this model supports additional user interactions on the basis of the creating results, where players are allowed to make further improvements on their profiles according to their needs.

III. SYSTEM STUDY

EXISTINGSYSTEM

The existing system consists of two categories: the model free methods and the model based methods.

A. *Model Based Methods*

Inthismethod,3Dfacesgeneratedarelimitedbytheamountofinformation contained in the model, they are usually missing in terms of details. There are three categories in model based methods:

- 3DMMfittingbasedmethod
- End-to-EnddeepCNNbasedmethod
- CascadedCNNbasedmethod

B. *Model free methods*

To avoid the limitations of model based methods researchers move on to this method. It uses volumetric CNN regression where it converts the whole complex network as regressed 3D output. Most of earlier methods are used to establish the correspondences of the special points between input images and 3D template including landmarks (i.e eye point, nose point, etc.,) and local features (i.e position and orientation of face images), then solve the nonlinear optimization function to regress the 3DMM coefficients.In these methods, even unsupervised methods have also been researched recently without a proper conclusion to lean upon. It can regress the 3DMM coefficients without the help of training data. Some other methods can reconstruct 3D faces without a 3D shape basis, it can produce a 3D structure by warping the shape of a reference 3D model.It also reconstructs the 3D shape of faces by learning a 3D Thin Plate Spline (TPS) warping function via a deep network which warps a generic 3D model to a subject specific 3D shape. In face alignment, 2D landmarks location only regresses visible points on faces, which is limited to describe face shape when the pose is large.After the 3D face reconstruction, Avatar rendering systems either require players to manually adjust considerable face attributes to obtain the desired face, or have limited freedom of facial shape and texture. These methods, generally use a GAN to imitate the game environment inorder to render the reconstructed face into an avatar. But these methods did not obtain desirable results for buxom and caricature faces.

IV. PROPOSED SYSTEM

An end-to-end method called PRN to jointly predict dense alignment and reconstruct 3D face shape is proposed. The proposed method is used to transfer the reconstructed 3DMM face shape to the game mesh which can be directly used in the game environment. This method specifies the representation and features of the position map. The CNN architecture and loss function are then used to learn mapping from an unconstrained RGB image to its 3D structure. From a single 2D image, regress the 3D facial geometry and its dense correspondence information.

As a result, a correct representation can be predicted directly using a UV position map with deep learning methods. A UV position map is a two-dimensional graph that records the three-dimensional positions of all points in UV space. UV space is used to record the 3D position of points from a 3D face model that is aligned with a 2D facial image. With the positive x-axis pointing to the right of the image and the minimum z at origin, the origin of the 3D space coincides with the upper-left of the input image. When projected to the x-y plane, the ground truth 3D facial shape completely matches the face in the 2D image.Using a CNN to regress the position map directly from unconstrained 2D images, the position map records a dense set of points from a 3D face with their semantic meaning in order to achieve the 3D facial structure and dense alignment result at the same time. The invisible parts of the face are also inferred by the position map therefore, this method can predict a complete 3D face.The 3D reconstructed face is given as an input to create an avatar. Differentiable rendering helps to design a rendering loop and force the 2D face rendering from the predicted shape and texture similar to the input reconstructed 3D face. This method consists of several trainable sub-networks. The image features are then flattened and fed to the lighting regressor which is a lightweight network consisting of several fully-connected layers and predicts lighting coefficients (light direction, ambient, diffuse and specular colour).Similar to image encoder, this method has a texture encoder. The features of the input image and coarse texture map are concatenated together and then fed into the texture decoder, producing the refined

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006082 |

texture map. After that the game mesh, the refined texture map, pose, and lighting coefficients are fed into the differentiable renderer to render the facemesh to a 2D image and enforce this image to be similar with the input 3D reconstructed face. To further improve the results, this method also introduces two discriminators, one for the rendered face image and another for the generated face texture maps.

V. MODULES

A. Data Processing

A strategy for creating a low-cost 3D face dataset is introduced. Thedataset used is 300W_LP. The 300W_LP dataset is a tensorflow official datasetwhich comes with classes like AFW, IBUG, HELEN, LFPW and 68 landmarksthat symbolises the important regions of the face while reconstructing it. Unlikeother approaches that necessitate single-view photos of subjects, which are notenough for accurate face reconstruction, this method necessitates multi-viewimages, which are enough for accurate face reconstruction. The flowchart ofdatasetprocessingisdepictedinFig(1).Thedataispreprocessedbyconvertingthe.matfilesto.npy.Aftergettingconvertedinto.npy(numpy)formattheUVcoordinatesareformedbytheshapeparameter and exponential parameters. With the angles of the input image from the post parameters along with the obtained transform parameters for the input data the UV position map isformed.

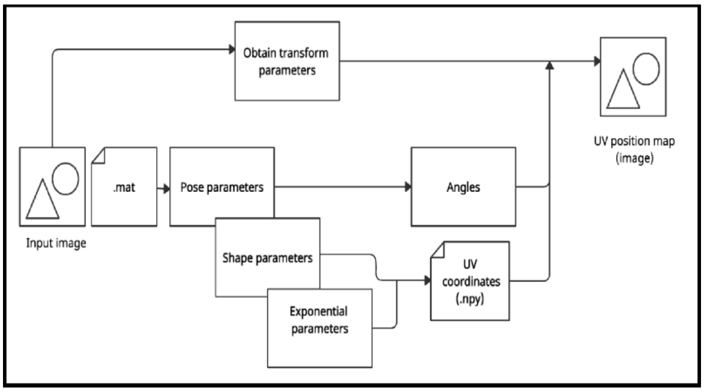


Fig. 1. Data processing – Block Diagram

B. 3D Face Reconstruction

300W_LP is used to create training sets since it contains face photos from various angles with annotations of predicted 3DMM coefficients, from which a 3D point cloud can be built quickly. The photos are cropped according to the ground truth bounding box and resized to 256 x 256 pixels. Then annotated 3DMM parameters are used to generate the appropriate 3D position and are rendered into UV space to obtain the ground truth position map; the map size in training is also 256 x 256, implying a precision of over 45K point cloud to regress. PRN is a simple self-supervised encoder decoder network which is trained using the prepared dataset. The position map, which is the output of the trained PRN is not limited to 3DMM face template or linear space. The target face is rotated and translated randomly on a 2D picture plane for the training.The rotation ranges from -45 to 45 degrees, the translation changes are random and range from 10% to 12% of the input size, and the scale ranges from0.9 to 1.2. It also enhances training data by scaling colour channels between 0.6 and 1.4. It synthesises occlusions by adding noise texture to raw photos in order to handle images with occlusions. Training data contains all of the challenging situations with all of the above augmentation operations.The flowchart of Face reconstruction is shown in Fig(2). Thus, a 3D face is reconstructed using the PRN as indicated.

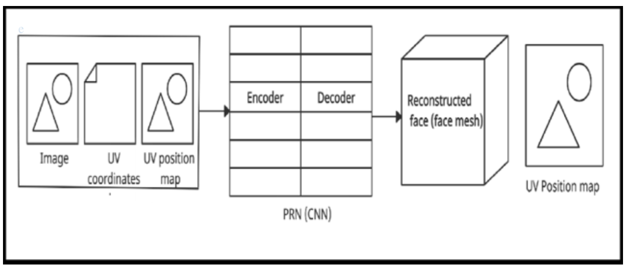


Fig. 2:3DFaceReconstruction-BlockDiagram

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006082 |

C. Avatar creation

The reconstructed face from the 3D face reconstruction module is given as an input here. The face mesh is passed to the shape transfer to generate a full head model instead of a front face only. Shape transfer converts the 3DMM mesh to a head mesh. The head mesh is also known as game mesh which is a collection of vertices, edges, and faces that make up a 3D object.The game mesh is converted into a texture map with the help of coarse texture. Coarse texture which is the facial structure of the face predicts the lighting and coefficients to refine the texture map. The texture map is a graphical design process.The texture decoder is used to remove all the unwanted features from the latent features which are obtained from the image encoder and texture encoder.Then the head mesh and the texture map is fed to the differential render along with a lighting coefficient which is predicted from the lighting predictor for the latent feature of the image in terms of pixels.The differential render is used to bridge the gap between 2D to 3D, minimising the difference between the rendered image and the input 2D image to get the desired avatar as displayed.

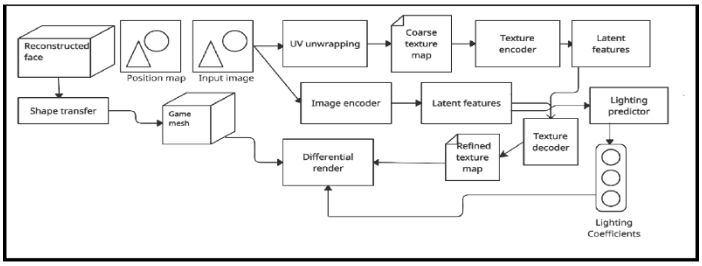


Fig3 : Avatar Creation -Block Diagram

D. User Interaction

The user interacts with the system by uploading their own image as a.jpg that is already taken or capturing a new image with the help of the system camera. The user's face is reconstructed in its full 3D form preserving its accuracy and landmark features including the texture and expressions as well. The flowchart is depicted in Fig(4) .The system is tested by uploading different images under different serendipity including caricature images and images with buxomfacesaswell

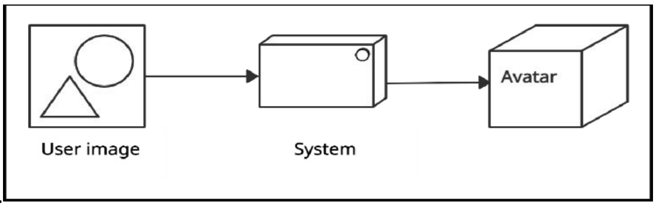


Fig 4:User Interaction-Block Diagram

VI. NETWORK -ARCHITECTURE DIAGRAM

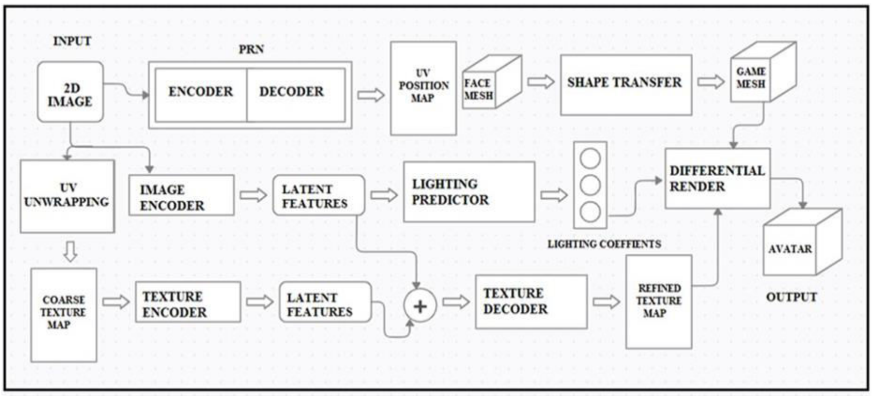


Fig 5 : Architecture Diagram

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006082 |

Initially 2D image of a human face has been fed into a simple PRN model which consists of encoder-decoder structure that takes image as an input and produces the output in the form of map or vector. Here the encoder consists of one convolutional layer and ten residual blocks. The convolutional layer extracts the features from an input image and also preserves the relationship between the input data by learning the features, the residual block is a stack of layers where the output of a layer is taken and added to another layer deeper in the block. Now the size of the image is reduced from 256 x 256 x 3 which is the original size to 8 x 8 x 512 which is the reduced size of the feature map in the encoder.

The output of the encoder is fed into a decoder which contains seventeen transposed convolutional layers to generate the predicted 256 x 256 x 3 position map by containing a modified input feature map. Kernel size is assigned as four for both convolutional and transposed convolutional layers. For the activation and to prevent the exponential growth in computation, a Rectified Linear Unit (ReLU) layer is used. Now finally the UV position map which holds the pixel of the image contains a complete 3D face structure which is known as face mesh.

Face Mesh is fed into a shape transfer module. Shape transfer module is used to transfer face mesh to game mesh. A face mesh is a 3D model of a face and a game mesh is a polygon mesh with the collection of vertices, edges and faces that make up a 3D object. For this process the shape transfer module is designed based on a radial basis function which is used to calculate distance between input and centre part of the image. The input is set as an original position of the mesh and the centre is represented as the original facial landmarks, using this parameter a radial basis function is calculated and a game mesh is created. Least square problem is used to minimise the distance between game mesh and face mesh to find out the best fit for the set of data points by minimising the sum of offset or residual points.

A 2D image of a human face is also fed into an image encoder which processes an image in terms of pixels and produces latent features which are the hidden features in terms of pixels for image. Once the latent features are extracted, it is passed into the lightning predictor which is used to predict the lightning coefficients. The lightning coefficients for the given latent features are predicted.Same 2D image of a human face is unwrapped in order to obtain the coarse texture map which holds the facial structure of the image in UV map format that is placing a 2D image in 3D plane. The texture map which is unwrapped from image is passed into a texture encoder which is used to process the image in terms of facial structure and produce the latent feature for the given texture map.

Once the latent feature of texture map is extracted, it is combined with a latent feature of the image which is processed from the image encoder and fed into the texture decoder which process the latent feature of the texture map and image to remove occlusion and produce an refine texture map which is clean and clear form of the facial image.The game mesh from shape transfer, the lightning coefficients from lightning predictor, the refined texture map from texture decoder are passed into the differential render which is designed to predict 3D object represented by mesh, point, map and thereby maximise the similarity between input and output image and it also trained to minimise the difference between input and output images. The outcome from the differential render is an avatar image of the given 2D input image which looks like a lifelike, non-photo realistic model in 3D form as displayed

VII. **CONCLUSION**

A simple low cost PRN based face reconstruction method is proposed which reconstructs faces for multiple views. The PRN is trained with a rich dataset that contains images with different poses, angles and occlusions, hence facilitating the PRN with the knowledge of reconstructing faces from images with diverse serendipity. This overcomes the difficulties faced by the previously proposed models which were not able to reconstruct faces with occlusions and different poses and angles, not able to consider the lighting of the images, neglected dense correspondence when model based methods were used and neglected UV space limit when model free methods were used. A differential render setting with texture encoder-decoder to preserve texture, lighting predictor to predict the lighting coefficients: is proposed which is able to create an avatar for the reconstructed face. It is able to create an avatar for both caricature images and images with buxom faces, which surpasses all the pre-existing methods that use GAN and imitator for avatar creation. The authenticity and tolerance of the system is checked by testing it with a varsity of images and the obtained outputs in all cases were found to be desirable.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006082 |

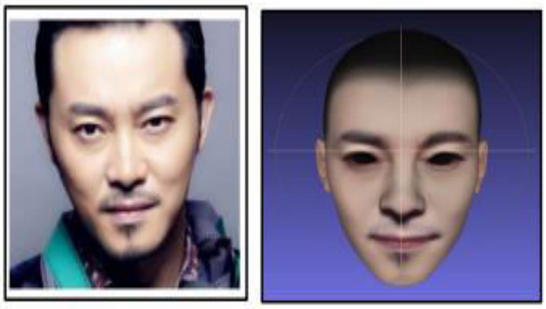


Fig 6: Avatar

VIII. FUTURE WORK

The avatar creation with face reconstruction and rendering system concentrates on reconstruction of face while fully preserving the features of the face in an accurate manner. It also focuses on reproducing the texture by predicting the lighting coefficients. The ability to negate the occluded region you texture over the real skin texture ie, the ability to revoke the occluded region texture while rendering the reconstructed face is yet to be implemented. Construction of a network that normalises texture evenly throughout the avatar by predicting the texture of occluded regions without being biased to the serendipity of the input picture, by training the network suitably is the future scope of the system. Avatar in itself is not used by users, they are generally used in a third party application after integrating with it. The third party applications may be

- Games: Games are highly effective which cause realism of the gaming experience to the players, when the first person player plays with an avatar.
- Animated movies: Animated Movies have a great effect if it has a character resembling an actor from the real world which helps the actor to attract their fans.
- Retail shopping: The avatars of actors in real time are used as an efficient marketing strategy. In addition to this, it attracts diverse customers from different ages, belonging to different ethnic groups. Contributes greatly to the advertisement purposes. Highly fit clothes are sold with the help of avatars.
- Telepresence: The still-with-you effect is provided by an avatar in meetings, conferences, personal calls and all such cases, even if the person could not make it to the venue in real time. Adds to the absence gap of the person.

REFERENCES

[1] Bas, Anil, Patrik Huber, William A. P. Smith, Muhammad Awais, and Josef Kittler. "3D Morphable Models as Spatial Transformer Networks." In 2017 IEEE International Conference on Computer Vision Workshops (ICCVW), 895–903. Venice, Italy: IEEE, 2017.https://doi.org/10.1109/ICCVW.2017.110.

[2] Blanz, V., and T. Vetter. "Face Recognition Based on Fitting a 3D Morphable Model." IEEE Transactions on Pattern Analysis and Machine Intelligence 25, no. 9 (September 2003): 1063-74. https://doi.org/10.1109/TPAMI.2003.1227983.

[3] Cheng, Yong, Zuoyong Li, Liangbao Jiao, Hong Lu, and Xuehong Cao. "Enhanced Retinal Modelling for Face Recognition and Facial Feature Point Detection under Complex Illumination Conditions." Journal of Electronic. Imaging 25, no. 4 (August 26, 2016): 043028. https://doi.org/10.1117/1.JEI.25.4.043028.

[4] Deng, Yu, Jiaolong Yang, Sicheng Xu, Dong Chen, Yunde Jia, and Xin Tong. "Accurate 3D Face Reconstruction With Weakly-Supervised Learning: From Single Image to Image Set." In 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 285–95. Long Beach, CA, USA: IEEE, 2019.https://doi.org/10.1109/CVPRW.2019.00038.

[5] Geng, Jiahao, Yanlin Weng, Lvdi Wang, and Kun Zhou. "Single-View Facial Reflectance Inference witha Differentiable Renderer."Science China Information Sciences 64, no. 11 (November 2021): 210101. https://doi.org/10.1007/s11432-020-3236-2.

[6] Jackson, Aaron S., Adrian Bulat, Vasileios Argyriou, and Georgios Tzimiropoulos. "Large Pose 3D Face Reconstruction from a Single Image via Direct Volumetric CNN Regression." In 2017 IEEE International Conference on Computer Vision (ICCV), 1031-39. Venice: IEEE, 2017. https://doi.org/10.1109/ICCV.2017.117.

[7] Lin, Jiangke, Yi Yuan, and Zhengxia Zou. "MeInGame: Create a Game Character Face from a Single Portrait." ArXiv:2102.02371 [Cs], February 6, 2021.http://arxiv.org/abs/2102.02371.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006082 |

[8] Liu, Shichen, Weikai Chen, Tianye Li, and Hao Li. “Soft Rasterizer: A Differentiable Renderer for Image-Based 3D Reasoning.” In 2019 IEEE/CVF International Conference on Computer Vision (ICCV), 7707–16. Seoul, Korea (South): IEEE, 2019.https://doi.org/10.1109/ICCV.2019.00780.

[9] Shi, Tianyang, Yi Yuan, Changjie Fan, Zhengxia Zou, Zhenwei Shi, and Yong Liu. “Face-to-Parameter Translation for Game Character Auto-Creation.” In 2019 IEEE/CVF International Conference on ComputerVision (ICCV), 161-70, Seoul, Korea (South): IEEE, 2019. https://doi.org/10.1109/ICCV.2019.00025.

[10] Shi, Tianyang, Zhengxia Zuo, Yi Yuan, Changjie Fan, Tianyang Shi, Zhengxia Zuo, Yi Yuan, and Changjie Fan. “Fast and Robust Face-to-Parameter Translation for Game Character Auto-Creation.” Proceedings of the AAAI Conference on Artificial Intelligence 34, no. 02 (April 3, 2020): 1733–40. https://doi.org/10.1609/aaai.v34i02.5537.

INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 8.165

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING