

The Avatar: 3-D Face reconstruction from two orthogonal pictures with application to facial makeover. (Final Report for B. Tech Project)

Nikhil Rasiwasia

Dept of Electrical Engineering, IIT Kanpur

nikux@iitk.ac.in

Under the Guidance of

Dr. K S Venkatesh, Dept of Electrical Engineering, IIT Kanpur

Abstract

In this report an algorithm for the fast reconstruction of a textured 3-D face model of the given individual from his two orthogonal pictures – a frontal view and a profile view - is presented. Then a possible application to the facial makeover of the individual is also conceptualized. The algorithm needs minimal human intervention for construction and does not need any special setup, or the camera calibration parameters as required by the Stereo based algorithm. Initially the facial features are identified and extracted giving the coordinated of the feature points, and then a generic model is deformed using Radial Basis Functions (RBF). The reconstructed model is in the standard virtual reality model (VRML) format, such that it can be viewed online by the common web browsers with a use of a plugin.

1. Introduction

3-D face reconstruction has been an active area of research in computer vision and graphics. The need for a fast reconstruction algorithm has emerged out in the recent past due to its application in realistic facial animation for low data video conferencing. Another important application is found in the gaming industry where there is a boom for realistic looking and feeling games. If the player can see himself as a part of the game system, the game gains popularity. Facial reconstruction has also been applied to the problem of face recognition. There is also a huge demand for such application in the movies and special effects industry where the realistic face is animated to obtain the desired effects. Later in this paper two more applications of this shall be introduced with the possibility of converting them to a business plan.

Generating the 3-D model is not an easy task. Various algorithms and abundant literature is available on this subject. The approaches for this problem are divided into three main categories [19]: pure image based rendering techniques, hybrid image based techniques and 3-D scanning techniques. In the first category of algorithms, the 3-D model is generated from the images only. They

do not try to estimate the real 3-D structure but, just interpolate between the given images. The Hybrid techniques uses some approximate information about the 3-D geometry and mixes it with the image based rendering to obtain more accurate results. The aim of both these algorithms is to obtain a coherent view of the real scene, not to obtain the metric measurement. In the third category of algorithms, complete 3-D structure is obtained. 3-D scanning techniques can be active – 3-D data from the range scanner or coded light, or can be passive. The passive methods are commonly known as Shape from X methods [22]. The commonly used passive methods are: Shape from Motion [20, 21], where one or multiple videos is used to find correspondences and then 3-D shape is extracted; Shape from Optical Flow, in which making correspondences is not important rather the apparent velocity of the pixels given by the optical flow field is used; Shape from Texture, which is a clear psycho-physical evidence of human use to extract depth. The perspective distortion and the size of the texel extracted from the low-level processing is used to find the 3-D shape; Shape from contour or Silhouette which aims to describe a 3-D shape as seen from one or more different directions [19]; Shape from Shading [23] based on diffusing properties of the Lambertian surfaces; Shape from Stereo, a widely used method for facial reconstruction [16, 17, 18, 19] that typically consists of three steps:- Camera Calibration, Establishing point correspondences between pairs of points from the left and the right image and Reconstruction of the 3-D co-ordinates of the points in the scene.

The algorithms mentioned above in this text suffer from two major problems. Firstly the quality of 3D reconstruction from Shape from X algorithms is often not satisfactory. Secondly, most of the algorithms require a special setup or a pre defined lighting to recover the 3d shape, making them unsuitable for the home environment. Many researches have successfully tried to tackle the first problem by using a generic model of the face, and deforming to fit the feature points obtained from the previous algorithms. For the second issue, two orthogonal views of the subject has been used by

researchers [2, 3, 4, 5] to simplify the process of obtaining the feature point information. If the two images are from arbitrary angle from the head, the corner point matching will fail as the skin may be smooth and free of blemishes. A frontal view and profile view of the subject has been used for the problem, which does not suffer from any of the issues mentioned above. We also use these images to obtain the co-ordinates of facial features in three dimensions and then deform a generic model by the use of Procrustes Analysis and Radial Basis functions.

2. Initialization

2.1. Geometric Model Choice

The human face has a basic structure that consists of features such as nose, mouth, eyes etc, for different people. A generic face model can easily encompass such features, but within these features there are differences that make one different from others. A generic model should be such that it can be modified to cover the space of all such different faces and should be structurally supportive of facial animation. The wireframe model can be modeled in softwares like 3D Studio Max, Maya, Poser etc.

This system uses a model that incorporates features like eyes, nose, mouth, forehead and chin and the in-between areas. The complete head model consisting of ears and hair has been deferred in this implementation. The facial model consists of 1683 3D vertexes and 3186 faces. (Fig. 2.1)



Figure 2.1- Generic Model

According to MPEG-4[24], a standard for which facial feature should be used in face animation has been set up. It defines 84 feature points on the neutral face called as the Facial Definition Parameters (FDPs) and/or Facial Animation Parameters (FAPs). 35 feature points are selected from the FDPs for the current system (Fig. 2.2). Another possible feature points model is the Candide-3 model [1], but consisting of 112 feature points.

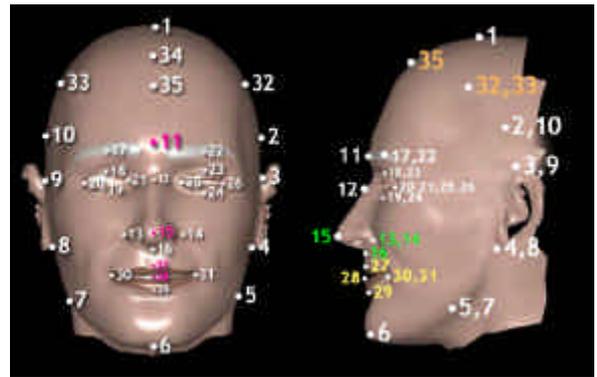


Figure 2.2 – Feature Points

This generic model is stored in virtual reality modeling language format also referred to as VRML97 [26]. The advantage of this format is that it can be viewed from different angles using commonly available plug-ins for the standard browsers.

2.2. Importing the Images

The frontal and the profile views can be obtained by a stereo camera setup as explained in [5], but for simplicity we allow even images from handheld camera. These images are aligned for processing as described in the next step. This relaxation does introduce an error as the images are assumed to be orthogonal, but they are perspective images in reality. A correction factor can be introduced in the later stages to minimize the error due to this assumption. Another important constrain is that the images are taken in normal white light condition on a background free from any skin colored object. The background may or may not be cluttered. (Fig 2.3)



Figure 2.3 – Input Images

2.3. Alignment of the profile view

If the images are not taken from a stereo camera setup, they need to be aligned by scaling, rotation and translation so that the features in the frontal and profile

view lie in the same horizontal line. The user is asked to mark 4 specific points – Nose, Eye, Ear and Mouth - in each image and then using these, the transformations are calculated.

$$q = \sin^{-1} \left(A / (B^2 + C^2) \right) - \tan^{-1} (C / B) \quad (2.1)$$

where theta is the angle by which profile image needs to be rotated. A = desired Y difference calculated from the ear and nose point from the frontal image, B and C are the actual X and Y difference between the ear and nose points obtained from the profile image. The scale factor is calculated from the difference in the Y co-ordinates of eye and mouth.



Figure 2.4 – Images after alignment

3. Image based Facial Feature Point extraction

The X, Y, Z co-ordinates of the 35 feature points are extracted from the frontal and the profile. The frontal view gives the X and Y coordinates - (x_f, y_f) . The profile view also gives the Y co-ordinate along with the Z co-ordinate - (z_p, y_p) . This Y co-ordinate is approximately same as y_f as a result of the alignment done in section 2.3. The final position of the feature points are:

$$(x_f[i], (y_f[i] + y_p[i]) / 2, z_p[i]) \text{ where } i = 1, 2, 3, \dots, 35. \quad (3.1)$$

In the current implementation 9 feature points – 4 of each eye and 1 for mouth - are extracted automatically as explained in section 3.2 and 3.3. For this the face region is identified in both the images. The remaining 26 features are found out by modifying a pre-determined set of 35 feature points – derived from the generic model – such that the squared distance between the corresponding 9 points extracted automatically is minimal.

3.1. Localization of the Face

Experience suggests that human face has a distinctive color that is generally not found in nature or common man made objects. Face localization based on color allows for fast processing (Fig 3.1). It is robust to geometric variations of the skin pattern and also robust under partial occlusion. This project employs a simple pixel based skin model that explicitly defines the skin region [25]. The model has been slightly modified to harden the conditions for a pixel to be classified as skin. It is based on the knowledge that in RGB space the skin has a higher Red content than other components. (R, G, B) is classified as skin if:

$$\begin{aligned} R > 95 \text{ and } G > 40 \text{ and } B > 20 \text{ and} \\ \max\{R, G, B\} - \min\{R, G, B\} > 15 \text{ and} \\ |R - G| > 15 \\ \text{and } R - G > 20 \text{ and } R - B > 20 \end{aligned} \quad (3.2)$$

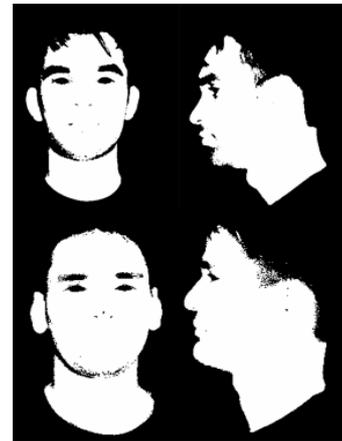


Figure 3.1 - Skin Detection

3.2. Extraction of the Eyes feature points from frontal image.

First step in determining the feature points is to localize the eye area (consisting of both the eyes) from the face. The Image of the subject is converted to binary levels and a template matching algorithm is applied [12, 13, 14]. Since template matching is a time intensive algorithm, the resolution of the face image is reduced to 64px width. The template (Fig. 3.2) is generic and captures the general properties of the eye and its surrounding area.

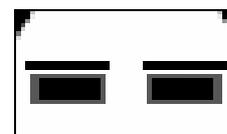




Figure 3.2 – Eyes Localized

Each eye region is extracted from the localized area and using the Prewitt operator the vertical edge detection is performed. The only vertical edges in and near the eye region belong to the eye itself (iris and eye boundary) and all other surrounding regions are either smooth or have strong horizontal components in their edges (eyebrows). The region with the maximum vertical edges gives the bounding box of the eyes. From this the X and the Y coordinates of the 4 points from each eye can be easily calculated. (Fig. 3.3)

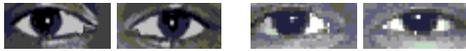


Figure 3.3 – Eyes Extracted

3.3. Extraction of the Mouth

Once the eye region is correctly identified, finding gross mouth region is trivial. A rectangular region is cropped from the face image with the extremities of the eyes as the right and left boundaries and the lower part of the eyes as the upper boundary. Horizontal histogram technique is implemented to find the location of the mouth in the image. For this we first apply a gray level morphological image reconstruction to improve the gray scale levels of the mouth and then convert the image to binary colors. The dark shadows and the background around the chin and neck interfere with the process. So such regions are tackled by removing the black connected components touching the image boundary. Location of the mouth is obtained by considering the first peak from the top above a certain threshold, in the horizontal histogram (Fig 3.4). The center of the mouth is determined by taking the vertical histogram in the binary image of the localized mouth region.

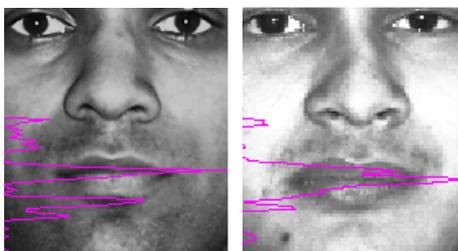


Figure 3.4 – Horizontal Histogram for mouth

3.4. Extraction of other features

The X,Y co-ordinates of 9 points are used to deform a pre-defined set 35 point derived from the generic model. The corresponding 9 points from the pre-defined model

are used to calculate values for scaling and translation in X and Y dimension and rotation. The problem can be formulated as minimizing the squared distance between the 9 model points after transformation and the observed points. In homogeneous co-ordinates:

$$\begin{bmatrix} x(i)_{mod} \\ y(i)_{mod} \\ 1 \end{bmatrix} = \begin{bmatrix} sx \times \cos q & -sx \times \sin q & tx \\ sy \times \sin q & sy \times \cos q & ty \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x(i) \\ y(i) \\ 1 \end{bmatrix} \quad (3.3)$$

Where $x(i)$, $y(i)$ are points from the pre-defined model
 $x(i)_{mod}$, $y(i)_{mod}$ are the points after transformations
 sx , sy : Scaling in X and Y dimension
 tx , ty : Translation in X and Y dimension
 $theta$: rotation in anticlockwise direction
and $i = 1, 2, \dots, 9$.

The function to be minimized is then as follows.

$$\min \{ f(sx, sy, tx, ty, q) \}$$

$$= \min \left\{ \sum_{i=1}^9 (x(i)_{mod} - a(i))^2 + (y(i)_{mod} - b(i))^2 \right\} \quad (3.4)$$

Where $x(i)_{mod}$ and $y(i)_{mod}$ is given by the Eq 3.3 and $(a(i), b(i))$ are coordinates of points as obtained from the frontal image.

This unconstrained non linear optimization problem can be solved using any of the commonly available methods such as Quasi-Newton Methods etc. (Fig 3.5)

For the Profile image, the tip of the nose is found out using the skin models developed in section 3.1. Using other heuristics a coarse estimate of the (z_p, y_p) can be obtained.



Figure 3.5– 35 features from Automatic methods

4. Interactive Facial Feature Point Editing

Even if all the 35 points are obtained using automatic methods, there might be some errors that creep in due

to the noise or ill-posed images. In this system since only a coarse estimate of the 26 points is obtained using automated methods the problem calls for some interactive technique to fine tune the co-ordinates. A GUI system had been developed in Matlab to facilitate the process of fine tuning. Initially the user is provided with the frontal image with the 35 feature points marked (Fig 4.1). These co-ordinates of these points can be changed by simply dragging them over the image. A prototype model is also displayed along with, which acts as a reference for the tuning process.

After the user is satisfied with the frontal image, a similar process is used for the profile image (Fig 4.1). The final Y co-ordinates of the feature points are then averaged over the two images. At the end of this all the information regarding the feature points is obtained. Now depending on the application, these can be further processed. If the use is in video conferencing, these values can be tracked over the video and transmitted to the other end for reconstruction. The reconstruction can also be done at the same end for other purposes.

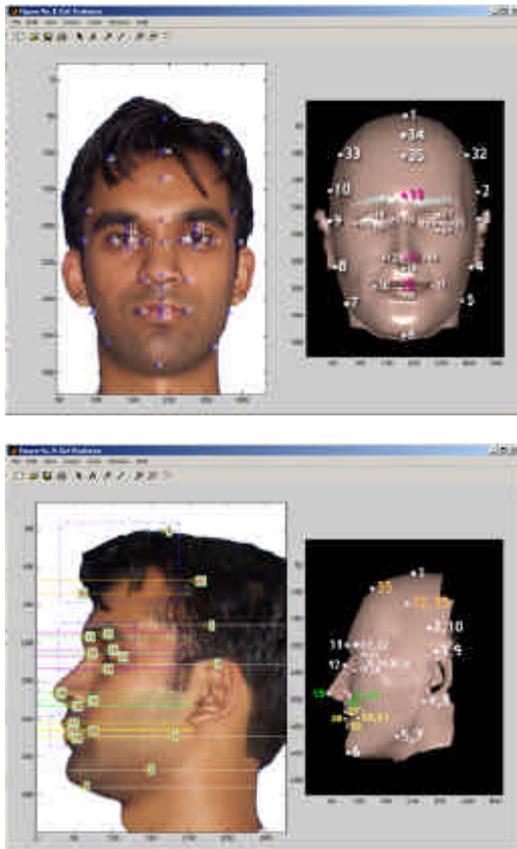


Figure 4.1 – GUI for interactive tuning

5. Generic Model Deformation

The generic model deformation is a problem about 3D space deformation in fact. A set of control points are identified in this space and their displacements are

computed. Then an interpolation function is chosen that accommodates the displacement of the control points. The other points in the 3-D space, in particular the points of the generic model are then calculated using the interpolation function.

We follow a two step process for deformation. First the global deformations is carried out in which the entire 3D model vertices is brought as close as possible to the feature points considering it to be a rigid body and in the second step the 3D model is treated as a non rigid object and points are matched in a more realistic manner, further closing on to the feature points.

5.1. Global Deformations

To achieve the goals of global deformation the 3-D model needs to be translated, scaled and rotated to match the feature points. This is a similar problem as tackled in section 3.4, except increase in dimensions from 2D to 3D and the number of points from 9 to 35. To solve this using the optimization approach is computationally expensive. A computationally simpler approach is to use the Procrustes Analysis [5, 27]. One drawback is that if the scale factor in X, Y and Z co-ordinates are not independent and attain the averaged values (Fig 5.1).

$$\min \{ f(s, R, T) \} = \min \left\{ \sum_{i=1}^{35} (P_{Model} - P_{Calculated})^2 \right\} \quad (5.1)$$

$$\text{Where } P_{Model} = \begin{bmatrix} x_m \\ y_m \\ z_m \end{bmatrix} = s \cdot R \begin{bmatrix} x_{m0} \\ y_{m0} \\ z_{m0} \end{bmatrix} + T \quad (5.2)$$

and s : Scale, R : Rotation matrix and T is a column vector for translation

$P_{calculated}$ is 35 x 3 matrix containing the feature points from the face images and P_{model} the corresponding points from the model.

- Compute the mean of both P_{model} and $P_{calculated}$
- Center each set at its origin (i.e. $P_{C0} = P_C - \text{mean}(P_C)$)
- Compute the norm of each set $\|P_{C0}\|^2$ and $\|P_{M0}\|^2$
- Normalize each set to equal unit norm (i.e. $P_{C0} = P_{C0} / \|P_{C0}\|^2$)
- Let $A = P_{C0}' P_{M0}$
- Compute the Singular Value Decomposition, SVD, of A which results in the matrices L, D , and M (i.e. $LDM = SVD(A)$)
- Compute the rotation matrix $R = M L'$

- h) Compute the scaling factor $s = \text{Tr}(D) * \frac{\|P_{C0}\|^2}{\|P_{M0}\|^2}$
- i) Compute the translating vector $T = \text{mean}(P_C) - s \cdot \text{mean}(P_M) \cdot R$
- j) Transform P_{Model} to $P_{\text{calculated}}$ using $P = s \cdot P_M \cdot R + T$
- k) Finally, transform the entire 3D model vertices using s , R and T .

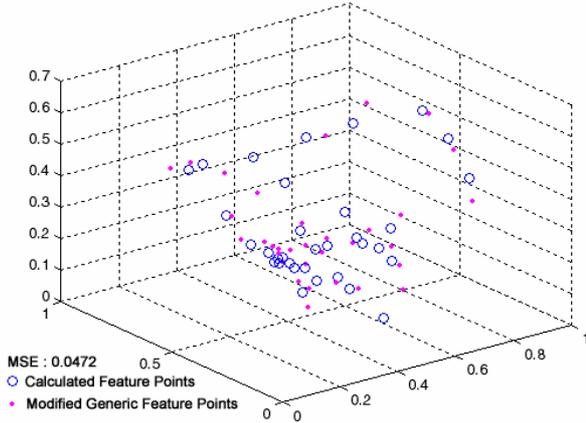


Figure 5.1 – Global Deformations

5.2. Local Deformations

In this second step a more realistic model is generated where the feature points from the generic model is actually mapped onto the feature points calculated. In this process an interpolation function is estimated by the displacements of the feature points and the rest of the points are mapped using this function. Constructing such an interpolation function is a standard problem in scattered data interpolation. Several techniques exist in the literature [9, 10]. An intuitive way to tackle the problem is by using distance weighted methods. Another approach is to construct the interpolant as a linear combination of basis functions and then to determine the coefficients of the basis functions.

$$f(p) = \sum_{i=1}^n w(i)R(\|p - p(i)\|) \quad (5.3)$$

where $w(i)$ are the weights to be determined using the feature points and the values of basis function R usually depends only on the distance from the feature points, and are thus called radial. Many different functions for $R(r)$ have been proposed [8, 10, 11]. We have chosen

$$R(\|p - p(i)\|) = e^{-\|p - p(i)\|/64} \quad (5.4)$$

We therefore find a smooth vector-valued function $f(P_{\text{Model}})$ fitted to the calculated feature points $P_{\text{calculated}}$ and

$$p_{x\text{Calculated}}^j = f_x(p_{x\text{Model}}^j) = \sum_{i=1}^{35} w_x(i)R(\|p_{x\text{Model}}^j - p(i)\|) \quad (5.5)$$

$j = 1, 2 \dots 35$

where $p_{x\text{Calculated}}^j$ is the X co-ordinate of the j th calculated feature point and $p_{x\text{Model}}^j$ gives the respective X co-ordinate from the generic model. Substitution of 35 feature points into equation 5.5 results in a linear system of 35 equation whose solution is of the form.

$$\begin{bmatrix} w_x(1) \\ w_x(2) \\ \vdots \\ w_x(35) \end{bmatrix} = [H^{-1}] \begin{bmatrix} p_{x\text{Calculated}}^1 \\ p_{x\text{Calculated}}^2 \\ \vdots \\ p_{x\text{Calculated}}^{35} \end{bmatrix} \quad (5.6)$$

The values for the weights for Y and Z co-ordinates are calculated similarly. Using the weights the mapping of other points can be trivially calculated. Upon experiment it was found that the points outside the convex hull of the feature points tend to get inside the hull considerably distorting the generic model. Eight new feature points were created to tackle these, whose co-ordinates were the vertices of the bounding cube of the generic model (Fig 5.2).

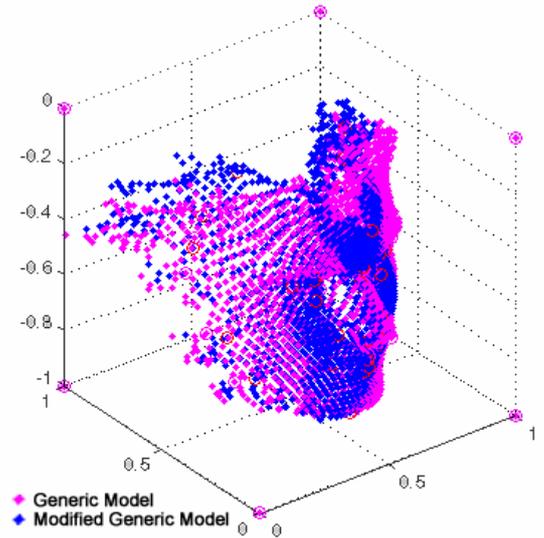


Figure 5.2 – Local Deformations

6. Texture Mapping

The final step in the reconstruction is texturing the modified 3-D model using the frontal image. We do not use the profile image for texturing as ears and hair is not a part of the generic model to be textured. With the generic model the texture co-ordinates are also saved in the vrml format. The job is to again deform the texture

co-ordinates so that the features overlap with the actual features in the frontal image. The procrustes analysis and the RBF can be again used in 2 Dimension for the transformation as the actual co-ordinates of the feature points are already known.



Figure 6.1 – Final Output 1

7. Facial Makeover

We also introduce two important applications of the facial reconstruction. First is for the online shopping industry and second in the High Tech Beauty Saloons. One of the major impeding factors of online shopping is helplessness of the customer as far as actually testing the product is concerned. For example a person wanting to buy sunglasses by the traditional ways of shopping would first try variegated sunglasses before making a purchase. Using the reconstructed model will enable the company to provide this feature online in a three dimensional environment where the customer will also be able to view his face from different angles not possible in the traditional way. Also companies such as L'Oreal which have quality associated with their brand are not able to sell online because the customer is not sure of which color will suit the best. This facility can also be implemented using image processing and image editing. For example the lips have already been extracted during the project, now the hue of the lips can be changed to match the color of the lipstick.

The second application in is the high end beauty saloons. People try different makeup solutions but are uncertainty of the results. It becomes difficult for them to visualize themselves in different appearances. This project will equip them with the power to test various different appearances possible and then take a decision. For example the user will actually be able to see in the how will he look with different hair model applied to his face.

8. Experimental Results

The complete system was developed in Matlab and was tested with imaged of Indian subjects including both the sexes (Fig 8.1, 8.2). The final textured model was viewed in Internet Explorer ver 6.0 using Cortona VRML Client 4.3 [28]



Figure 8.1 – Final output 2



Figure 8.2 – Female image

9. Conclusion and Future work

In this paper we have described an image based approach to reconstruct the 3D model from two orthogonal images – Frontal and Profile view. A generic model has been successfully deformed to generate the individualized model of the subject. A sub process of this required calculating the feature points on the face, currently 9 of which are successfully calculated automatically. Finally the texture is mapped on to the model to make it more realistic.

The emphasis on this paper has been on developing the proof of concept with minimizing the user aid and maximizing the quality of the output. Possible ways to improve the quality and reduce the user intervention are:-

- Using a higher resolution model of the face with more points and vertices.
- Taking the pictures using a stereo camera setup instead of handheld cameras
- A more robust algorithm for face extraction based on Gaussian models can be implemented.
- More feature points can be automatically located, in fact the number of feature points itself can be increased

10. Acknowledgement

We are thankful to the people who volunteered to be a part of this project and allowed their pictures to be used in the project.

11. References

- [1] J. Ahlberg, "CANDIDE-3, an updated parameterized face,". Technical Report No. LiTH-ISY-R-2326, Dept. of Electrical. Engineering, Linköping University
- [2] WS Lee, NM Thalmann, "Fast head modelling for animation", Journal Image and Vision Computing, 18(4), pp355-364, 2000.
- [3] Lee W, Kalra P, Magnenat Thalmann N (1997), "Model Based Face Reconstruction for Animation", MMM '97, Singapore, Nov 1997
- [4] Mandum Zhang, Linna Ma, "Image Based 3D Face Modelling", IEEE Computer Society
- [5] A-Nasser Ansari, Mohamed Abdel-Mottaleb: 3-D Face Modeling Using Two Views and a Generic Face Model with Application to 3-D Face Recognition. AVSS 2003: 37-44
- [6] S. Kshirsagar, S. Garchery, N. Magnenat-Thalmann: "Feature Point Based Mesh Deformation Applied to MPEG-4 Facial Animation" , Proceedings Deform'2000
- [7] 3D face modeling using two orthogonal views and a generic face model Ansari, AN Abdel-Mottaleb, M. Dept. of Electr. & Computer Engineering, ICME '03
- [8] N. Arad, N. Dyn, et al., "Image Warping by Radial Basis Functions: Application to Facial Expressions," Graphical Models and Image Processing, Vol 56 1994
- [9] D. Ruprecht, R. Nagel, and H. Müller: "Spatial Free-Form Deformation with Scattered Data Interpolation Methods", Computers & Graphics 19(1), 1995, pp.63-71.
- [10] Jun-yong Noh, Douglas Fidaleo, Ulrich Neumann: "Animated deformations with radial basis functions". VRST 2000: 166-174
- [11] Pighin, F., Hecker, J., Lischinski D., Szeliski, R., and Salesin, D."Synthesizing Realistic Facial Expressions from Photographs" ACM SIGGRAPH'98.
- [12] Hua Gu Guangda Su Cheng Du, "Feature Points Extraction from Faces" Research Institute of Image and Graphics, Department of Electronic Engineering, Tsinghua University, Beijing, China
- [13] Karin Sobottka and Ioannis Pitas, "Face Localization and Facial Feature Extraction Based On Shape And Color Information," Proceedings of the IEEE 1996
- [14] R. Lanzarotti, NA Borghese, and P. Compadelli, "Robust Identification and Matching of Fiducial Points for the. Reconstruction of 3D Human Faces from Raw Video Sequences," Proceedings of the First International. Symposium on 3D Data Processing Visualization and Transmission' 02
- [15] Chung J. Kuo, Ruey-Song Huang, Tsang-Gang Lin: 3-D facial model estimation from single front-view facial image. IEEE Trans. Circuits Systems for video Technology, vol 12 March 2002
- [16] Jie Zou, Peng-Jui Ku, Luyun Chen, "3-D Face Reconstruction Using Passive Stereo"
- [17] R. Enciso, J. Li, D. Fidaleo, TY. Kim, JY.Noh, U. Neumann, "Synthesis of 3D Faces", International Workshop on. Digital and Computational Video, 2000
- [18] P. Fua , C. Miccio, "Animated heads from ordinary images: a least-squares approach", Computer Vision and Image Understanding, v.75 n.3, p.247-259
- [19] Esteban, C. E., Schmitt,F.: "Silhouette and Stereo Fusion for 3D Object Modeling". Fourth International Conference on 3-D Digital Imaging and Modeling , pp. 46–54 October 06-10, 2003, Banff, Alberta, Canada.
- [20] C.-M. Cheng and S.-H. Lai, "An Integrated Approach to 3D Face Model Reconstruction from Video", IEEE Proc. Of the Second International Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems (RATFG-RTS 2001), pp.16-23, Vancouver, Canada, July 2001
- [21] AR Chowdhury, R. Chellappa, S. Krishnamurthy, and T. Vo, "3D Face Reconstruction from Video using a. Generic Model," IEEE International Conference on Multimedia and Expo, 2002. ICME '02. Volume: 1 , 26-29 Aug
- [22] "Image Processing Analysis, and Machine Vision." Pg 508-558, Milan Sonka, Vaclav Hlavac, Roger Boyle.
- [23] Mohamad Ivan Fanany, Itsuo Kumazawa: "Analysis of shape from shading algorithms for fast and realistic 3-D face recognition." APCCAS 2002
- [24] <http://www.chiariglione.org/mpeg/standards/mpeg-4/mpeg-4.htm>, International Organisation for Standardisation - Coding of Moving Pictures and Audio
- [25] Vezhnevets V., Sazonov V., Andreeva A., "A Survey on Pixel-Based Skin Color Detection Techniques". Proc. Graphicon-2003, pp. 85-92, Moscow, Russia.
- [26] "<http://www.web3d.org/x3d/specifications/vrml/ISO-IEC-14772-IS-VRML97WithAmendment1/>" Information technology -- Computer graphics and image processing -- The Virtual Reality Modeling Language (VRML)
- [27] RM Everson. "Orthogonal, but not orthonormal, procrustes problems." In Advances in Computational Mathematics, 1998
- [28] <http://www.parallelgraphics.com/products/cortona/> - VRML Client

