

RESEARCH PAPER

Available Online at www.jgrcs.info

3D FACE RECOGNITION AND MODELING SYSTEM

Sushma Jaiswal^{*1}, Dr. Sarita Singh Bhadauria², Dr. Rakesh Singh Jadon³

¹Lecturer, S.O.S. in Computer Science, Pt. Ravishankar Shukla University, Raipur (C.G.)
jaiswal1302@gmail.com

²Professor & Head, Department of Electronics Engineering, Madhav Institute of Technology & Science, Gwalior (M.P.)

³Professor & Head, Department of Computer Applications, Madhav Institute of Technology & Science, Gwalior (M.P.)

Abstract: In this paper, 2D photographs image divided into two parts; one part is front view (x, y) and side view (y, z). Necessary condition of this method is that position or coordinate of both images should be equal. We combine both images according to the coordinate then we will get 3D Models (x, y, z) but this 3D model is not accurate in size or shape. In defining other words, we will get 3D face model, refinement of 3D face through edit of point and smoothing process. Smoothing is performed to get the more realistic 3D face model for the person. We measure to compare the average time for modeling and compare the research result of our methods with different techniques, for this purpose we taken by two hypotheses (1) the average quality of our method will be higher than the 60% (2) it is faster compare to other in an average case (3) it is automated. First hypothesis is correct but the second tie up with other three methods and third found satisfactory.

Keywords-3D model, face recognition, hypothesis, 2D photographs, Smoothing.

INTRODUCTION AND LITERATURE REVIEW

Face recognition has been an active research area over last 40 years. This research spans several disciplines such as image processing, pattern recognition, computer vision, and neural networks. It has been studied by scientists from different areas of psychophysical sciences and those from different areas of computer sciences. Psychologists and neuroscientists mainly deal with the human perception part of the topic, whereas engineers studying on machine recognition of human faces deal with the computational aspects of face recognition. Face recognition has applications mainly in the fields of biometrics, access control, law enforcement, and security and surveillance systems.

The problem of face recognition can be stated as follows: Given still images or video of a scene, identifying one or more persons in the scene by using a stored database of faces [R.Chellappa et.al.(1995)]. The problem is mainly a classification problem. Training the face recognition system with images from the known individuals and classifying the newly coming test images into one of the classes is the main aspect of the face recognition systems.

This problem seems to be easily solved by humans where limited memory can be The main problem; whereas the problems for a machine face recognition system are:

1. Facial expression change
2. Illumination change
3. Aging
4. Pose change
5. Scaling factor (i.e. size of the image)
6. Frontal vs. profile
7. Presence and absence of spectacles, beard, mustache etc.
8. Occlusion due to scarf, mask or obstacles in front.

The problem of automatic face recognition (AFR) is a composite task that involves detection of faces from a cluttered background, facial feature extraction, and face identification. A complete face recognition system has to solve all sub problems, where each one is a separate research problem. This research work concentrates on the problem of 3D face Recognition and modeling.

Many researches in face recognition have been dealing with the challenge of the great variability in head pose, lighting intensity and direction, facial expression, and aging. The main purpose of this overview is to describe the recent 3D face recognition algorithms. The last few years more and more 2D face recognition algorithms are improved and tested on less than perfect images. However, 3D models hold more information of the face, like surface information, that can be used for face recognition or subject discrimination. Another major advantage is that 3D face recognition is pose invariant. A disadvantage of most presented 3D face recognition methods is that they still treat the human face as a rigid object. This means that the methods aren't capable of handling facial expressions. Although 2D face recognition still seems to outperform the 3D face recognition methods, it is expected that this will change in the near future.

Chua et al. [1997, 2000] introduced point signatures to describe the 3D landmark. They used point signatures to describe the forehead, nose and eyes. Their method reached a recognition rate of 100% when tested on a dataset with 6 subjects. Wang et al. used the point signatures to describe local points on a face (landmarks). They tested their method on a dataset of 50 subjects and compared their results with the Gabor wavelet approach [Wiskott, L. et.al.(1997)]. Their results showed that point signatures alone reached a recognition rate of 85% where the Gabor wavelets reached a recognition rate of 87%. If both 2D and 3D landmarks were

combined, they reached a recognition rate of 89%. The authors remarked that these results could also be influenced by the number of landmarks used for face recognition, since for the point signatures 4 landmarks were used, for the Gabor wavelets 6 landmarks and for the combination of both 12 landmarks were used.

Douros and Buxton proposed the Gaussian Curvature to define quadratic patches to extract significant areas of the body. They claim that their method can be used for recognition of all kinds of 3D models [Douros, I., Buxton, B.F.(2002)]. Another local shape descriptor that was found to perform good on human bodies was the Paquet Shape Descriptor [Robinette, K.M.(2004)].

Blanz, Vetter and Romdhani proposed to use a 3D morphable model for face recognition on 2D images [Blanz, V et.al.(2002,2003)– Romdhani, S., Vetter, T.(2003)]. However, the recognition rate was for all approaches of the morphable model between the 75% and the 99%.

Naftel et al. presented a method for automatically detecting landmarks in 3D models by using a stereo camera [Huang, J et.al.(2003)]. The landmarks were found on the 2D images by an ASM model. These landmark points were transformed to the 3D model by the stereo camera algorithm. This algorithm was correct in 80% of all cases when tested on a dataset of 25 subjects.

A similar idea was proposed by Ansari and Abdel-Mottaleb [2003]. They used the CANDIDE-3 model [Ahlberg, J.(2001)] for face recognition. Based on a stereo images landmark points around the eyes, nose and mouth were extracted from the 2D images and converted to 3D landmark points. A 3D model was created by transforming the CANDIDE-3 generic face to match the landmark points. The eyes, nose and mouth of the 3D model were separately matched during the face recognition. Their method achieved a recognition rate of 96.2% using a database of 26 subjects.

Suikerbuik [2004] proposed to use Gaussian curvatures to find 5 landmarks in a 3D model. He could find the correct landmark point with a maximal error of 4 mm . Gordon proposed to use the Gaussian and mean curvature combined with depth maps to extract the regions of the eyes and the nose. He matched these regions to each other and reached a recognition rate of 97% on a dataset of 24 subjects [Gordon, G.G.(1991)]. Moreno et al. used both median and Gaussian curvature for the selection of 35 features in the face describing the nose and eye region [Moreno, A.B. et.al.(2003)]. The best recognition rate was reached on neutral faces with a recognition rate of 78%. Xu et al. proposed to use Gaussian-Hermite moments as local descriptors combined with a global mesh [Xu, C. et.al.(2004)]. Their approach reached a recognition rate of 92% when tested on a dataset of 30 subjects. When the dataset was increased to 120 subjects, the recognition rate decreased to 69%.

Lu et al. had used the generic head from Terzopoulos and Waters [Terzopoulos, D., Waters, K.(1993)] which they adapted for each subject based on manually placed feature points in the facial image [Lu, X.et.al.(2004)]. Afterwards the models were matched based on PCA. This method was

tested on frontal images and returns in 97% of all cases the correct face within the best 5 matches. 3DMeNow Pro v.2 by BioVirtua human modeling package. Used to build recognizable human and character models for interactive 3D games and broadcast animation, this software produces amazingly lifelike 3D head data (models, textures and morph states) in a fraction of the time taken with conventional manual authoring techniques.

Faceworx by looxis application is a free software which creates a 3D head out of two standard 2D photos. You need two well illuminated pictures; one from the front (mugshot style) and one from the side. The software demands some skills in placing reference points and marking the contours of the face; mouth, nose, ears and eyes. The final 3D portrait can be saved and exported in the well known OBJ format.

CyberExtruder's AvMaker software will automatically create a 3D head model of the person in the picture--there are no points to move around and no need to locate the eyes--the software does it all! The speed with which AvMaker generates a new head model is truly impressive because of how quickly and accurately it locates all the anatomically relevant parts of the human head in a 2D image resulting in a very fast 2D to 3D generator.

Haar-like features have scalar values that represent differences in average intensities between two rectangular regions. They capture the intensity gradient at different locations, spatial frequencies and directions by changing the position, size, shape and arrangement of rectangular regions exhaustively according to the base resolution of the detector. Appearance-based methods Most of these algorithms use raw pixel values as features. However, they are sensitive to addition of noise and change in illumination. Instead, Papageorgiou et al. [1998] used Haar-like features, which are similar to Haar basis functions. The features encode differences in average intensities between two rectangular regions, and they can extract texture without depending on absolute intensities.

Jaiswal et.al.(2010) gives Brief Description of literature on Image Based human and machine recognition of faces during 1987 to 2010. Machine recognition of faces has several applications. As one of the most successful applications of image analysis and understanding, face recognition has recently received significant attention, especially during the past several years. In addition, relevant topics such as Brief studies, system evaluation, and issues of illumination and pose variation are covered. In this paper numerous method which related to image based 3D face recognition are discussed.

Recently, Viola and Jones proposed an efficient scheme for calculating these features [P. Viola and M. Jones (2001)]. They also proposed a method for constructing a strong classifier by selecting a small number of distinctive features using AdaBoost. This framework provides both robustness and computational efficiency.

Jaiswal et. al (2008) describes an efficient method and algorithm to make individual faces for animation from possible inputs. Proposed algorithm reconstruct 3D facial model for animation from two projected pictures taken from

front and side views or from range data obtained from any available resources. It is based on extracting features on a face in automatic way and modifying a generic model with detected feature points with conic section and pixelization. Then the fine modifications follow if range data is available. The reconstructed 3D face can be animated immediately with given parameters. Several faces by one methodology applied to different input data to get a final Animatable face are illustrated.

Many improvements or extensions of this method have been proposed. We will introduce two major approaches and then point out problems of these approaches. The first approach is an improvement of the boosting algorithm. There are modified versions of AdaBoost such as Real Boost [R. E. Schapire and Y. Singer (1999)], KLBoosting [C. Liu and H. Y. Shum (2003)] and FloatBoost [S. Z. Li et.al.(2004)]. The second approach is using extensions of the feature sets such that various image patterns can be evaluated. As shown in Figure 1, in addition to the basic feature set (a), different arrangements or numbers of rectangles such as (b), (c) or (d) are used in [P. Viola and M. Jones(2001), B. Wu et.al.(2004)].

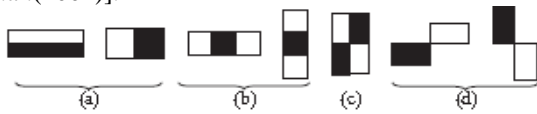


Figure 1. Examples of Haar-like feature sets.

Lienhart et al. [2002] introduced an efficient scheme for calculating 45° rotated features. Although both approaches are effective, they are insufficient to achieve more accurate face detection. Most of these methods construct a weak classifier by selecting one feature from the given feature set. However, the generalization performance is no longer improved in later rounds of the boosting process because the classification task using only one feature becomes more difficult. Viola and Jones reported that features which were selected in later rounds yielded error rates between 0.4 and 0.5 while features selected in early rounds had error rates between 0.1 and 0.3 [P. Viola and M. Jones (2001)].

Jaiswal et.al.(2007) “Automatic 3D Face Model from 2D Image-Through Projection “ take 2D images and convert it into 3D model this model recognize the face. In this method get 3D animatable face, refinement of 3D animatable face through pixelization and smoothing process. Smoothing is performed to get the more realistic 3D face model for the person.

OUR METHOD -3D FACE RECOGNITION AND MODELING SYSTEM

In this section we describe our algorithm, which is capable of automatically or semi-automatically construct the 3D face model from frontal and profile face. The process of our algorithm can be divided into the following steps:

Front View and Side View:

Both views of face are captured in such a manner that the horizontal distance of front view and side view are same. The images should be in the same size (i.e., same pixel size). During capturing the side view care should be taken that the

side view is perpendicular to the horizontal line front view. As the input images to parameters detection could be used photographs in an optional resolution and quality, but we recommend to use images with resolution 512x512 and higher, where more information about facial features is included and high resolution texture could be created here we taken the Reference of -the extended XM2VTS database.(see Figure 2)



(a)

Choose left side of the face or Right side of the face



(b)

(c)

Figure2 (a) Shows Front view of Face (b) Left Side view of Face (c) Right Side View of Face(XM2VTS dabse)

Face Extraction:

The face is extract from the front view in such a manner that the vertical dimension of the face from the front view and side view should be same and horizontal dimension of side view should cover the face. The human face and its properties detection should be performed separately for both images. To recognize the head and gather its properties from images two main techniques are used. Haar Cascade Classifiers based on an extended set of Haar like features.

Getting Coordinate through Projection of Front View (x, y):

In this step, x coordinates are taken from the extracted image of front view taking reference of left or right border of the image and y coordinate are taken from bottom border of image. Each pixel of the front view can be identified by the set of x and y coordinates.(see Figure 3)



Figure3 Getting coordinate through projection of front view (x, y)[

XM2VTS dabse]

Getting Coordinates through Projection of Side View (y, z):

In this step, z coordinates are taken from the extracted image of side view taking reference of right border of the image and y coordinate are taken from bottom border of image. Each pixel of the side view can be identified by the set of y and z coordinates. To detect parameters from the profile image we apply a median filter to smooth the segmentation result and reduce the noise. (see Figure 4)

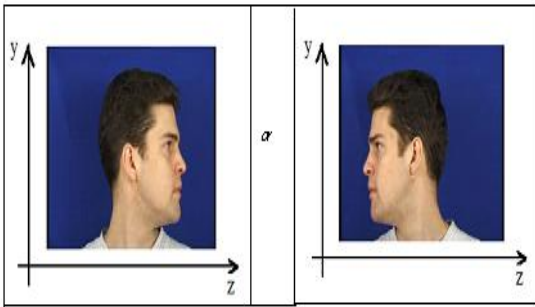


Figure 4 Getting coordinates through projection of side view (y, z) [XM2VTS database]

3D Modeling (x, y, z):

3D coordinates taken from previous method now in the three dimensional spaces. This three dimensional plot of x, y and z coordinates will show the 3D face model. Here one thing is important that the y coordinate of both views Y_{Front} and Y_{Side} are the same or close to each other they can be averaged of the y axis for each know corresponding 2D in a combined 3D(x,y,z) in the relative image plane coordinate. Detecting face through the with use Haar Cascade Classifiers and image based segmentation. Here 3D face geometry construction by transforming the predefined model according to the parameters detected in this step and Color texture and normal map texture creation and mapping (see Figure5 and Figure6).



Figure 5 Creation of texture through YCrCb color space



Figure 6 Normal map

Smoothing the 3D Face Model:

After the Creation of 3D the image must be filtered so that for a given pixel and every neighboring pixel are compared to the 3D original image and the difference

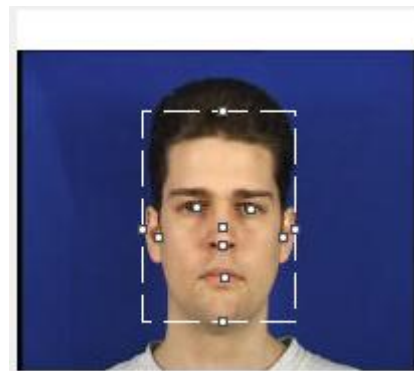
between these pixels are used to smoothen the image through correction of points. We apply the depiction in 3D models so real time 3D model were created. Geometry of the Face model is represented as triangle mesh and vertices are coded for each triangle separately. Therefore remove doubles option should be used in a digital content creation tool to reach smooth and continuous surface. The image encoding is converted into YCrCb color space and for all pixels the Cr and Cb values are evaluated. YCrCb is an encoded nonlinear RGB signal, Color is represented by luma (which is luminance, computed from nonlinear RGB [Poynton 1995]), constructed as a weighted sum of the RGB values, and two color difference values Cr and Cb that are formed by subtracting luma from RGB red and blue components. (see figure 7)

$$Y = 0.299R + 0.587G + 0.114B$$

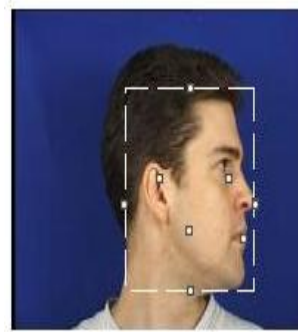
$$Cr = R - Y$$

$$Cb = B - Y$$

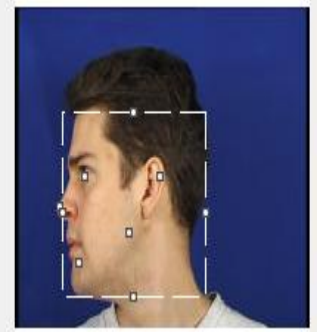
The transformation simplicity and explicit separation of luminance and chrominance components makes this colorspace attractive for skin color modelling [Phung et al. 2002].



(a)



(b)



(c)

Figure 7 (a) front image (b) left side image (c) Right Side Image, Correction of points [XM2VTS database]

3D Face Model:

Color texture is automatically created by joining detected face areas from front and profile input images. The first step is scaling a smaller image up to have same face height as in the bigger one. Then a front face image rectangle from frontal image is joined with a profile face image rectangle. A profile rectangle is connected to the front rectangle from both left and right sides. Horizontally flipped profile image is joined from the left side. X joint coordinate is in front image left eye left border position and right eye right position and eye center in profile image. Blending between images is calculated by the linear interpolation in

connection area. After image junction, known texture parts are cloned into the background texture parts. For example the hair texture is cloned into the space above head. When the color texture creation is finished, we can calculate a normal map texture from it. Grayscale copy of color texture is used as a surface height map, because it contains main head surface details information. For each pixel in grayscale image it is calculated its x direction and y direction intensity difference with the next pixel in current direction. This difference is scaled by surface height scale constant to decrease differences. We found empirically the height scale constant 0.01 in our experiments. X and Y differences are brought into account as derivatives of the image function in X and Y directions. From these derivatives we can create directional vectors and calculate image function normal as their cross product. Then we normalize obtained normal and transform its coordinates into the RGB color space for storing in the normal map texture. Finally it will give the 3D face model.

Environment

In order to develop and test the solution, code must be written. Two possibilities have been identified for the development environment to be used; they are as follows:

- C++ (or another such compiled language).
- VB6.0

While a compiled language is potentially quicker at runtime and more universal (which would benefit a more commercial application), the scope of this project is to provide a proof of concept. It is for this reason that C++ was chosen, due to a number of key features. Being a scripted language, it allows faster development and rapid prototyping of new code, at the possible expense of slower runtime. Many required functions are built in to the C++ development environment, removing the need for these to be adapted to the compiled code. A production system could be transferred to a compiled language, when development and testing cycles are not as dominant.

RESULTS AND CONCLUSION

We tested developed 3D Face modeling system on Intel(R) core™2 Duo CPU 2.66GHz, 2047MB RAM computer with NVIDIA GeForce 7300 SE/7200 GS graphic card. Detection in a front image takes 1.00 seconds in average and detection in a profile image takes 1.00 seconds in average in this hardware configuration. The 3D Face model updating process takes 5 seconds on average. 3D Face modeling system accepts the common image data formats supported by Photograph of front and side as the input. For example jpg, bmp, png, tif file formats are supported. Human Face from front or profile view should be in input photographs. We recommend image resolution 440x440 or higher for better face recognition and more details in created color texture. The main advantage of Haar cascade classifiers is the robust and precise face and facial feature detection. So we use the respective methods for our modeling system. The quality of construction was subjectively evaluated for three approaches by 20 randomly selected peoples in age from 24 to 50 years. The selection of peoples was not dependent from age, sex and education.

Peoples were watched during construction evaluation. We compare our methods with other methods (1) 3dMenowpro

(2) 3d morphable face (3) cyber extruder based model (4) Faceworx. We examined construction quality of our approach was subjectively ranked to 60% is reported. So the first hypothesis is proved(see in Table 1) but the other techniques are same or better in average time so the second hypothesis is false because the other two methods gives the same average time i.e. 6 seconds(see Table 2). The average time is affected by the processor speeds or high configuration of processor. In future we will try to modify the structure for second hypothesis.

It is hard to compare the results of different methods to each other since the experiments presented in literature are mostly performed under different conditions on different sized datasets. For example one method was tested on neutral frontal images and had a high recognition rate, while another method was tested on noisy images with different facial expressions or head poses and had a low error rate.

Some authors presented combinations of different approaches for a face recognition method and these performed all a little better than the separate methods. But besides recognition rate, the error rate and computational costs are important, too. If the error rate decreases significantly, while the recognition rate increases only a little bit, the combined method is still preferred. But, if the computational costs increase a lot, calculation times could become prohibitive for practical applications.

Table 1 comparison of methods in average time

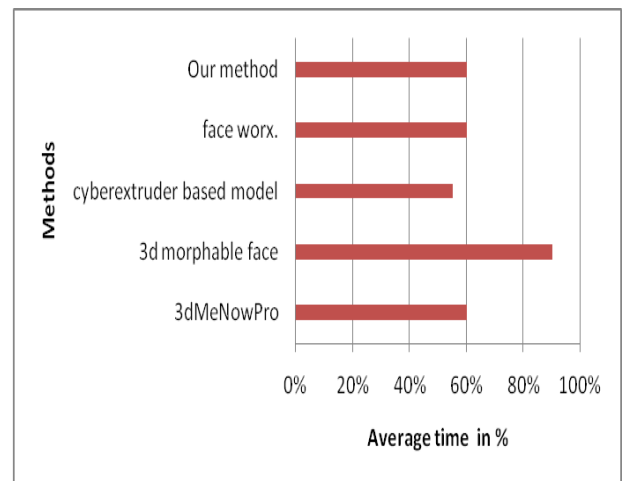
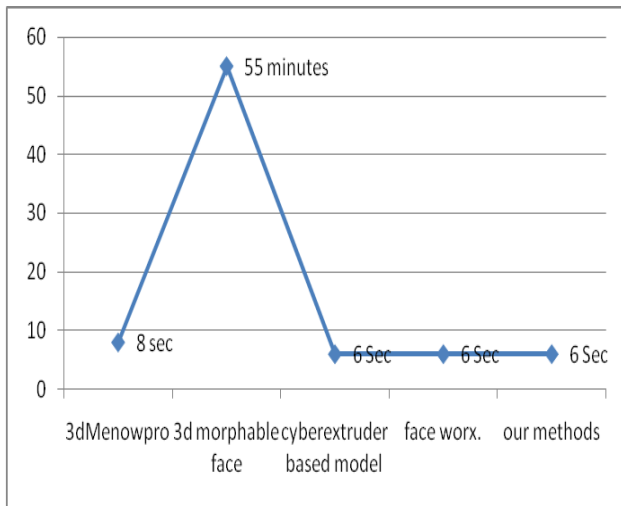


Table 2 Comparison of methods according time taken



REFERENCES

- [1] Suikerbuik, C.A.M., Tangelder, J.W.H., Daanen, H.A.M., Oudenhuijzen, A.J.K.: Automatic feature detection in 3D human body scans. Proceedings of the conference "SAE Digital Human Modelling for Design and Engineering. (2004)
- [2] Gordon, G.G.: Face Recognition Based on DepthMaps and Surface Curvature. Proceedings of the SPIE, Geometric Methods in Computer Vision, Vol. 1570. (1991) 108–110
- [3] Moreno, A.B., Sánchez, A., Velez, J.F., Díaz, F.J.: Face Recognition using 3D Surface-Extracted Descriptors. Proceedings of the Irish Machine Vision and Image Processing Conference. (2003)
- [4] Jaiswal, S., S.S. Bhaduria and R.S. Jadon, 2007. Automatic 3D face model from 2D image-through projection. Inform. Technol. J., 6: 1075-1079.
- [5] Jaiswal, S., S.S. Bhaduria and R.S. Jadon, 2008. Creation 3D Animatable Face Methodology Using Conic Section-Algorithm, Inform. Technol. J., 7, 292-298.
- [6] Xu, C., Wang, Y., Tan, T., Quan, L.: Automatic 3D Face recognition combining global geometric features with local shape variation information. Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition. (2004) 302–307
- [7] Chua, C.S., Jarvis, R.: Point Signatures: A New Representation for 3D Object Recognition. International Journal on Computer Vision. 25 (1997) 63–85
- [8] Chua, C.S., Han, F., Ho, Y.K.: 3D human face recognition using point signature. Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition. (2000) 233–238.
- [9] Sushma Jaiswal, Sarita Singh Bhaduria, Rakesh Singh Jadon and Tarun Kumar Divakar(2010), Brief description of image based 3D face recognition methods, springer,3D Research Volume 1, Number 4, 1-15, DOI: 10.1007/3DRes.04(2010)02.
- [10] Wiskott, L., Fellous, J.M., Kruger, N., van der Malsburg, C.: Face Recognition by Elastic Bunch Graph Matching. IEEE Transactions on Pattern Analysis and Machine Intelligence. 19 (1997) 775–779
- [11] Douros, I., Buxton, B.F.: Three-Dimensional Surface Curvature Estimation using Quadric Surface Patches. Proceedings of the Scanning 2002 Conference. (2002)
- [12] Robinette, K.M., An Alternative 3D descriptor for database mining. Proceedings of the Digital Human Modelling Conference. (2004)
- [13] Blanz, V., Romdhani, S., Vetter, T.: Face Identification across different poses and illuminations with a 3D morphable model. Proceedings of the IEEE International Automatic Face and Gesture Recognition. (2002)
- [14] Blanz, V., Vetter, T.: Face Recognition Based on Fitting a 3D Morphable Model. IEEE Transactions on Pattern Analysis and Machine Intelligence. 25 (2003)
- [15] Romdhani, S., Vetter, T.: Efficient, Robust and Accurate Fitting of a 3D Morphable Model. Proceedings of the European Conference on Computer Vision. (2003)
- [16] Heisele, B., Koshizen, T.: Components for Face Recognition. Proceedings of the Audio- and Video-Based Biometric Person Authentication, Vol 2688. (2003) 153–159
- [17] Huang, J., Heisele, B., Blanz, V.: Component-Based Face Recognition with 3D Morphable Models. Proceedings of the Audio- and Video-Based Biometric Person Authentication, Vol 2688. (2003) 27–34
- [18] Naftal, A.J., Mao, Z., Trenouth, M.J.: Stereo-assisted landmark detection for the analysis of 3D facial shape changes. Technical Report TRS-2002-007. Department of Computation UMIST, Manchester. (2002)
- [19] Ansari, A., Abdel-Mottaleb, M.: 3D Face modeling using two views and a generic face model with application to 3D face recognition. Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance. (2003) 37–44
- [20] Ahlberg, J.: CANDIDE-3 - an updated parameterized face. Technical Report LiTHISY- R-2326. Dept. of Electrical Engineering, Linköping University. (2001)
- [21] Terzopoulos, D., Waters, K.: Analysis and Synthesis of Facial Image Sequences Using Physical and Anatomical Models. IEEE Transactions on Pattern Analysis and Machine Intelligence. 15 (1993) 569–579
- [22] Lu, X., Hsu, R., Jain, A., Kamgar-Parsi, B.: Face Recognition with 3D Model-Based Synthesis. Proceedings of the International Conference on Biometric Authentication. (2004) 139–146.
- [23] biovirtual-3dmenow- professional. software. informer.com
- [24] <http://www.looxis.com>.
- [25] www.cybextruder.com.
- [26] C. P. Papageorgiou, M. Oren, and T. Poggio. A general framework for object detection. *Proc. of ICCV*, pages 555–562, 1998.
- [27] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. *Proc. of CVPR*, pages 511–518, 2001.
- [28] R. E. Schapire and Y. Singer. Improved boosting algorithms using confidence-rated predictions. *Machine Learning*, 37(3):297–336, 1999.
- [29] C. Liu and H. Y. Shum. Kullback-leibler boosting. *Proc. Of CVPR*, pages 587–594, 2003.

- [30] S. Z. Li and Z. Q. Zhang. Floatboost learning and statistical face detection. *IEEE Trans. on PAMI*, 26(9):1112–1123, 2004.
- [31] B. Wu, H. Ai, C. Huang, and S. Lao. Fast rotation invariant multi-view face detection based on Real AdaBoost. *Proc. of IEEE Conf. on Automatic Face and Gesture Recognition*, pages 79–84, 2004.
- [32] R. Lienhart and J. Maydt. An extended set of haar-like features for rapid object detection. *Proc. of ICIP*, 1:900–903, 2002.
- [33] Reference-the extended XM2VTS database, <http://www.ee.surrey.ac.uk/CVSSP/xm2v>

