International Journal For Technological Research In Engineering
Volume 2, Issue 10, June-2015 *ISSN (Online): 2347 - 4718*

PERFORMANCE EVALUATION OF DIFFERENT 3D FACE RECOGNITION TECHNIQUES WITH VARIATIONS IN EXPRESSION

Yuvaraj G. Melge[1], Satish P. Deshpande[2], Abhishek V. Deshmukh[3]
[1]M.Tech 4th SEM, [2,3]Asst Professor
EC Dept, KLS GIT, Belgavi, Karnataka.

Abstract: The variations in the facial expressions are the most critical sources in face recognition; especially in the frequent case where for enrollment per person single sample is available. For an reliable authentication system the methods that improves the accuracy in such variations are still required. In our project this problem is analyzed using synthesis based scheme and number of synthetic face images are produced with different expressions. For each user 3D animatable model is generated based on the automatically located 17 landmark points. Using these additional images the recognition performance is evaluated with two different techniques, which are by using principal component analysis (PCA) and linear discriminant analysis (LDA), on both face recognition grand challenge (FRGC) and Bosphorus 3D face databases. For each database and algorithm significant improvements in face recognition accuracies are achieved.
Keywords: Face recognition, Facial expression, Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA).

I. INTRODUCTION

The recognition of the human is very crucial in the security applications and grows continuously. Biometry enables the efficient identify management system by exploring the physical and behavioral characteristics of the subject that should be easy to access. Each have their own limitation and numerous biometric system to detect various human characteristics like iris, voice, face, fingerprint among those the traits is not realistic concept . There is a common assumption that though of the same person the face shape may change, due to various facial expressions, still there are regions which will keep their position and shape or be subjected to much less deformation among different expressions. If these regions can be identified, a 3D non-rigid face recognition problem can be reduced to the rigid case. Our fundamental purpose is to provide recognition by simulating facial expressions on 3D models of every single subject. With regard to the causes of the intra-class variations, synthesis of facial images under various pose and illumination conditions using 3D face models is straightforward since these variations are external. However, this will not be a positive indication for those expressions that change the facial surface characteristics in addition to appearance. For achieving realistic facial expression simulations, we represent an automated procedure for generating MPEG-4 compliant animatable face models from the 2.5D facial scans (range images) of the enrolled subjects based on a set of automatically detected feature points. Employing an engine on the name of facial animation, we can facilitate every single person with different expressions and further the images that are synthesized are adopted as additional gallery samples for the purpose of task recognition. On a mandatory stress note synthetic sample augmentation will be done during the curse of enrolment only once for every single subject.

II. OVERVIEW

Recognition of humans has become a substantial topic today as the need for security applications grows continuously. Biometry enables reliable and efficient identity management systems by exploiting physical and behavioral characteristics of the subjects which are permanent, universal and easy to access. The motivation to improve the security systems based on single or multiple biometric traits rather than passwords and tokens emanates from the fact that controlling a person’s identity is less precarious than controlling what he/she possesses or knows. Additionally, biometry-based procedures obviate the need to remember a PIN number or carry a badge. Each having their own limitations, numerous biometric systems exist that utilize various human characteristics such as iris, voice, face, fingerprint, gait or DNA. “Superiority” among those traits is not a realistic concept when it is parted from the application scenario. The system constraints and requirements should be taken into account as well as the purposes of use-context that include technical, social and ethical factors. For instance, while fingerprint is the most wide-spread biometric trait from a commercial point of view (mainly due to a long history in forensics), it mostly requires user collaboration. Similarly, iris recognition, which is very accurate, highly depends on the image quality and also requires significant cooperation from the subjects.

III. SYSTEM OVERVIEW AND ARCHITECTURE

In the proposed system, the enrollment is assumed to be done in both 2D and 3D for each subject under a controlled environment – frontal face images with a neutral expression and under ambient illumination. The obtained 3D shape of the facial surface together with the registered texture is preprocessed, firstly to extract the face region. On the extracted facial surface, scanner-induced holes and spikes are cleaned and abi lateral smoothing filter is employed to remove white noise while preserving the edges.

After the hole and noise free face model (texture and shape) is obtained, 17 feature points are automatically detected using either shape, texture or both, according to the regional properties of the face. These detected points are then utilized to warp a generic animatable face model so that it completely transforms into the target face. The generic model with manually labeled 71 MPEG-4 points is suitable to simulate facial actions and expressions via an animation engine that is in accordance with MPEG-4 Face and Body Animation (FBA) specifications. Finally, in order to simulate the facial expressions on the obtained animatable model, an animation engine, called visage|life™1 is utilized. Multiple expression-infused face images are generated for each subject to enhance face recognition performance. The whole system is illustrated in Fig. 3.1.

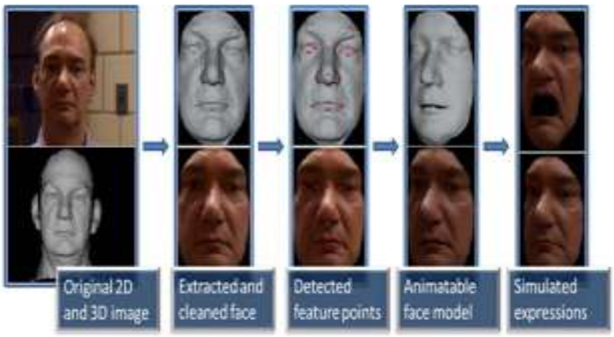

Fig. 3.1 The enrollment procedure is illustrated on an example.

IV. ARCHITECTURE DESIGN

Decomposition Description

A. Data Pre-processing

3D scanner outputs are mostly noisy. The purposes of the preprocessing step can be listed as:

A. To extract the face region (same in 2D and 3D images);
B. To eliminate spikes/holes introduced by the sensor;
C. To smooth the 3D surface.

Firstly, the nose tip is detected: For each row, the position with the maximum z value is found and then for each column, the number of these positions is counted to create a histogram. The peak of this histogram is chosen as the column for the position of the vertical midline, and the maximum point of this contour is identified as the nose tip. Using a sphere of radius 80mm and centered 10mm away from the nose tip in+z direction, the facial surface is cropped. Next, the existing spikes are removed by thresholding. Spikes are frequent with laser scanners, especially in the eye region. After the vertices that are detected as spikes are deleted, they leave holes on the surface. Together with other already existing holes (again usually around the eyes and eyebrows), they are filled by applying linear interpolation. Once the complete surface is obtained, a bilateral smoothing filter is employed to remove white noise while preserving the edges. This way, the facial surface is smoothed but the details hidden in high frequency components are maintained.

B. Automatic Land marking

Bearing in mind that subject cooperation is required during the enrollment, we base our system on the assumption of a well-controlled acquisition environment in which subjects are registered with frontal and neutral face images. In accordance with our scenario, we aim to extract a subset (17 points) of MPEG-4 Facial Definition Parameters (FDPs) to be utilized for the alignment of the faces with the animatable generic model. For the extraction of the points, 2D and/or 3D data are used according to the distinctive information they carry in that particular facial region. Firstly, facial midline (vertical profile) analysis is done and 5 fiducially points on that midline are detected. Based on that information; face is split into sub-regions for the coarse localization of eyes, nose and lips. After that, further analysis is done inside these extracted sub-regions to detect the points of interest. For those regions with non-informative texture (like nose), 3D data is analyzed. On the other hand for the regions with noisy surface and/or distinctive color information (like eyes), 2D data is utilized. As a result, 17 facial interest points are detected in total, consisting of 4 points for each eye, 5 points for the nose and 4 points for the lips (Fig. 4.1).

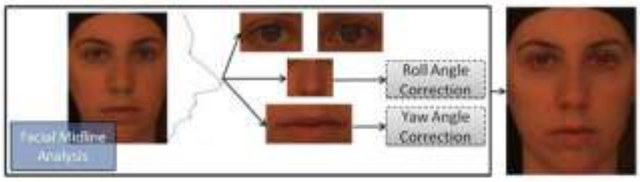

Fig. 4.1. The outline of the landmarking process and the detected feature points as the outcome.

V. RESULT AND CALCULATIONS

A. Vertical Profile Analysis:

The analysis done on the vertical profile constitutes the backbone of the whole system. It starts with the extraction of the facial midline and for this purpose; the nose tip is detected as explained previously. The nose tip position allows us to search for the eyes in the upper half of the face in order to approximately locate irises, so that the roll angle of the face can be corrected before any further processing. For coarse iris extraction, the non-skin region is found by removing the pixels with the most frequent chrominance values present in the upper half of the face in YCbCr space. Edge maps are constructed for the non-skin region using Canny edge detector by iteratively adjusting the threshold until a descriptive map is obtained. Subsequently, Hough transform is applied to the edge map to detect circles. For each detected circle, an overlapping score is calculated by the ratio of the detected portion of the circle to the whole circle parameter. After the detected iris candidates are grouped as right and left according to their relative positions to the profile line, the one with the maximum total overlapping score is selected among the compatible pairs. Next, the 2D and 3D images of the face are rotated in order to align the detected iris centers on the same horizontal line. Thereby, our assumption for vertical profile is better assured.

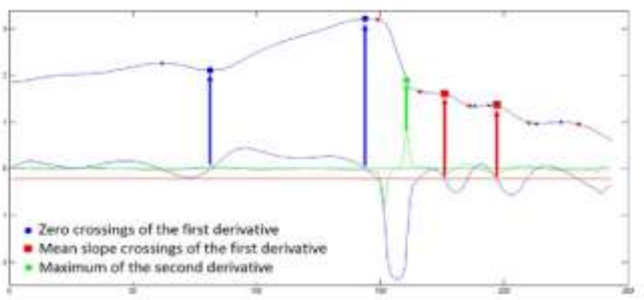

Fig. 5.1. A sample profile curve and its first (blue) and second (green) derivative curves. The arrows show the five detected interest points among the candidates.

B. Eye Regions:
The 3D surface around the eyes tends to be noisy because of the reflective properties of the sclera, the pupil and the eyelashes. On the other hand, its texturecarries highly descriptive information about the shape of the eye. For that reason, 2D data is preferred and utilized to detect the points of interest around the eyes, namely the iris center, the inner and outer eye corners and the upper and the lower borders of the iris. After applying an averaging filter with a rectangular kernel and the noise and the horizontal edges are suppressed, the vertical edges are detected with a Sobel operator. Using the vertical edge image, the irises are once again detected by Hough transform, this time more accurately. For the detection of eye corners, horizontal edges that belong to the eyelids are detected as described in and two polynomials are fitted for lower and upper eyelids. The inner (near the nose bridge) and outer eye corners are determined as the intersection points of the two fitted polynomials. After this step, the image is rotated one last time in the 2D image plane, if necessary, to horizontally align the two iris centers (Fig. 5.2).

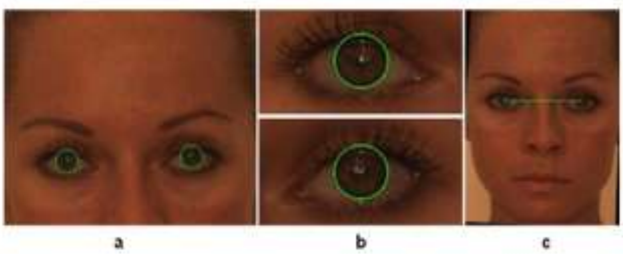

Fig.5.2 (a) Initial detection of iris circles in the upper half of the face. (b) Result of the detailed analysis for each eye. (c) Re-rotated face to align iris centers.

C. Nose Region:
Contrary to the eye region, nose region is extremely distinctive in surface but quite plain in texture. For this reason, we choose to proceed in 3D. To start with, the yaw angle of the face is corrected in 3D. For this purpose, the horizontal curve passing through the nose tip is examined. Ideally, the area under this curve should be equally separated by a vertical line passing through its maximum (assuming the nose is symmetrical). With this assumption, the curve is iteratively rotated to minimize the difference between these two partitions under the curve. Once the angle is determined, the whole surface is rotated, so that a "more frontal" face is obtained. After this adjustment, the minimum curvature is calculated for each point in the nose region. Then, edge detection is applied on the minimum curvature map. Those edges reveal the position of the points of interest on both sides of the nose tip (Fig. 5.3). The other three points; on the nose bridge, on the nose tip and at the bottom of the nose are already found during the vertical profile analysis.

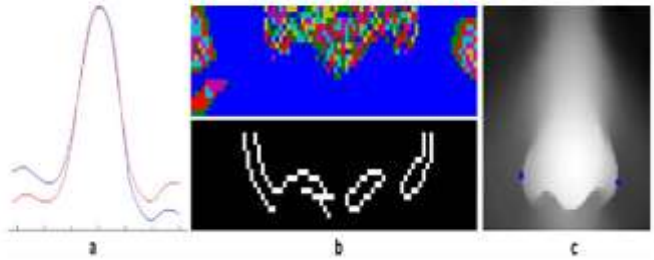

Fig.5.3 (a) The horizontal profile of the nose tip before (blue) and after corr ection (red). (b) The minimum curvature and the corresponding edge map. (c) The depth map of the nose region with detected points marked.

D. Lips Region:
Numerous studies are proposed to extract the lip contour based on deformable or active contour models working only with 2D images. However with our system, a much simpler method can be easily adopted because we have good estimates of two points of interest (upper and lower mid points) on the lip, thanks to the analysis performed on the vertical midline of the face. Since we work on faces with neutral expressions, the mouth is assumed to be closed. A closed mouth always yields to a darker line between the two lips. Based on this knowledge, the contact point of the lips is found by applying a vertical projection analysis. Once the 17 points are obtained, they are used to align the generic face to the target face for the construction of the animatable model.

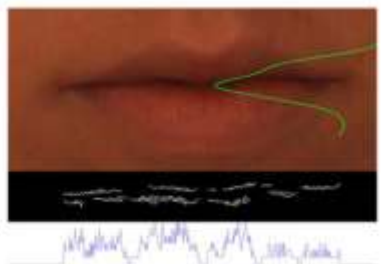

Fig. 5.4. A lip image is given with the detected points. The green line shows the horizontal projection result. At the bottom, the calculated edge map and its vertical projection is given.

VI. COMPONENT DESIGN

Because of the facial muscle contractions, soft tissue of the face deform during expression variations, which affects the recognition performance. In this section we review several 3-D face-recognition approaches and they are classified into three main categories based to handle expression variations. For a facial soft tissue deformations a statistical model is produced which is caused by expressions using a training dataset of expression builds by some methods. We called these methods as a statistical method. And other methods because of the variations in the expressions to be isometric assume the deformation, means that along the length of the surface the deformation is preserved. We refer these methods as the isometric deformation modeling approaches. The region-based methods which is third class of method does not assume a isometric deformation or statistical model but it uses only regions which are not or not much affected by the expressions. These three classes are in more detail are discussed in the following sections. The typical process is as shown in Fig. 6.1.

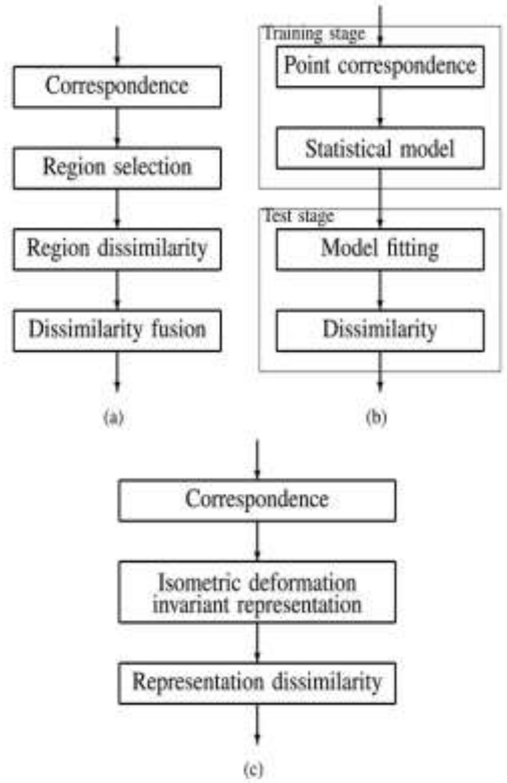

Fig.6.1. (a) Typical region-based 3-D face-recognition algorithm structure. (b) Algorithm using a statistical model, and (c) Iisometric deformation model.

VII. REQUIREMENT SPECIFICATION

Software Requirement
Operating system: Windows XP/7/Vista.
Coding Language: MATLAB

Hardware Requirement
Processor: Dual Core
RAM: 2GB
Mouse
Keyboard
HDD: 500GB

VIII. CONCLUSION

The implantation of image processing and face recognition using MATLAB is practically proved. Based on the assumption of a fully-controlled environment for enrollment, a face recognition framework is proposed in which the widely-encountered single sample problem for identification of faces with expressions is targeted by augmenting the dataset with synthesized images. A animatable model is created on automatic detection of 17 landmarks is presented utilizing 2D facial data. The experiments are conducted on large and well accepted database FRGC. The experiment results reveal that introduction of realistically synthesized face images with expressions improves the performance of the identification system. The face recognition performance is evaluated using PCA and LDA algorithm and the result shows that LDA gives more accurate result than PCA.

REFERENCES

[1] J. P. Phillips et al., "FRVT 2006 and ICE 2006 large-scale experimental results, "IEEE Trans, Pattern Anal. Mach. Intell, vol. 32, no. 5, pp. 831–846, May 2009.
[2] Y. Hu, D. Jiang, S. Yan, L. Zhang, and H. Zhang, "Automatic 3D reconstruction for face recognition," in Proc. IEEE Int. Conf. Autom. Face Gesture Recognit., May 2004, pp. 843–848.
[3] Lu and A. K. Jain, "Matching 2.5D face scans to 3D models, "IEEE Trans, Pattern Anal Mach. Intell., vol. 28, no. 1, pp. 31–43, Jan. 2006.
[4] F. Tsalakanidou, S. Malassiotis, and M. G. Strintzis, "Integration of 2D and 3D images for enhanced face authentication," in Proc. IEEE Int. Conf. Autom. Face Gesture Recognit., May 2004, pp. 266–271.

