3-D Face Recognition with Facial Expressions

Poojali Koti
M.Tech Student
Department of Computer Science and Engineering
Siddharth Institute of Engg & Tech.

J. Suneetha
Associate Professor
Department of Computer Science and Engineering
Siddharth Institute of Engg & Tech.

Abstract:— Face Recognition is one of the most suitable applications of image analysis. Taking so many sources in face recognition based on that one of the most critical sources is facial expressions. Human emotional facial expressions play an important in interpersonal relations. This paper presents a face recognition done with an analysis by synthesis based scheme. In which a number of synthetic face images with different expression are produced by using Data Preprocessing, Automatic Land marking and Animatable Face Models. Face Recognition performance done with linear decrement analysis principal component analysis, Local binary patterns techniques and Bosphorus 3D face databases.

Keywords— Face recognition, facial expressions, 3d facial animations

I. INTRODUCTION

Various government agencies are now more motivated to improve security data systems based on body or behavioral Characteristics, often called biometrics (Perronnin and Dugelay, 2003). In general, biometric systems process raw data in order to extract a template which is easier to process and store, but carries most of the information needed know rather than which a person really is. Biometry enables reliable and efficient identity management systems by exploiting physical and behavioural 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. For instance, while fingerprint is the most widespread biometric trait from a commercial point of view (mainly due to a long history in forensics), it mostly require user collaboration. Similarly, iris recognition, which is very accurate, highly depends on the image quality and also requires significant cooperation from the subjects. Face recognition stands out with its favorable reconciliation between accessibility and reliability Face recognition system can help in many ways: for example some applications are Checking for criminal records and Detection of a criminal at public place, Finding lost children's by using the images received from the cameras fitted at some public places and detection of thief’s at ATM machines, Knowing in advance if some unknown person is entering at the border checkpoints and so on. A face recognition system can operate in either or both of two modes: (1) face verification (or authentication), and (2) face identification (or recognition). Face verification involves a one-to-one match that compares a query face image against a template face image. Face identification involves one-to-many matches that compare a query face image against all the template images in the database to determine the identity of the query face. The first automatic face recognition system was developed by Kanade [2], so the performance of face recognition systems has improved significantly.

Face Recognition Vendor Test 2002 (Phillips et al., 2002), the vast majority of face recognition methods based on 2D image processing using intensity or color images, reached a recognition rate higher than 90% under lighting controlled conditions, and whenever subjects are consentient. Unfortunately in case of pose, illumination and expression variations the system performances drop, because 2D face recognition methods still encounter difficulties. In a recent work, Xu et al. (2004) compared intensity images against depth images with respect to the discriminating power of recognizing people. From their experiments, the authors concluded that depth maps give a more robust face representation, because intensity images are heavily affected by changes in illumination. Generally, for 3D face recognition is intended a class of methods that work on a three-dimensional dataset, representing both face and head shape as range data or polygonal meshes. The main advantage of the 3D based approaches is that the 3D model retains all the information about the face geometry. Moreover, 3D face recognition also growths to be a further evolution of 2D recognition problem, because a more accurate representation of the facial features leads to a potentially higher discriminating power. In a 3D face model, facial features are represented by local and global curvatures that can be considered as the real signature identifying persons.
Like other biometric traits, the face recognition problem can also be briefly interpreted as identification or verification of one or more persons by matching the extracted patterns from a 2D or 3D still image or a video with the templates previously stored in a database. Despite the fact that face recognition has been drawing a never-ending interest for decades and major advances were achieved, the intra-class variation problems due to various factors in real-world scenarios such as illumination, pose, expression, occlusion and age still remain a challenge. As 3D sensing technologies advance and the acquisition devices become more accurate and less expensive, the utilization of range data instead of / together with the texture data broadens.

The main advantage of the 3D based approaches is that the 3D model retains all the information about the face geometry. Moreover, 3D face recognition also grows to be a further evolution of 2D recognition problem, because a more accurate representation of the facial features leads to a potentially higher discriminating power. In a 3D face model, facial features are represented by local and global curvatures that can be considered as the real signature identifying persons.

In this paper, address the problem of expressions in 2D probe images. Our aim is to facilitate recognition by simulating facial expressions on 3D models of each 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 does not hold for expressions which alter the facial surface characteristics in addition to appearance. In order to achieve realistic facial expression simulations, we propose an automatic procedure to generate 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. Using a facial animation engine, different expressions are simulated for each person and the synthesized images are used as additional gallery samples for the recognition task. It is important to emphasize that synthetic sample augmentation is carried out during enrolment only once for each subject. Creating various 2D synthetic faces could be a good way to overcome the classical problems of 2D face recognition. First of all, we have to consider that modern 3D computer graphics technologies are able to reproduce synthetic images in an excellent realistic way and with an accurate geometric precision. Secondly, we have to consider that 3D facial reconstruction from a single view image can be considered good enough, only if the experimental results show a high discriminating power.

II. FACE RECOGNITION ARCHITECTURE

In the above figure 1 Overview of our complete face recognition system on an FPGA. Video data is received from the camera and sent to the face detection subsystem which finds the location of the face(s). These face(s) can be any size. The architecture then performs down sampling of the detected face to $20 \times 20$, and sends these 400 pixel values to the recognition subsystem.
Face recognition is a challenging research area in terms of both software (developing algorithmic solutions) and hardware (creating physical implementations). A number of face recognition algorithms have been developed in the past decades with various hardware implementations. All previous hardware implementations assume that the input to the face recognition system is an unknown face image. Current hardware based face recognition systems are limited since they fail if the input is not a face image. A practical face recognition system should not require the input to be a face, instead would recognize face(s) from any arbitrary video which may or may not contain face(s) potentially in the presence of other objects. Therefore, an ideal face recognition system should first have a face detection subsystem which is necessary for finding a face in an arbitrary frame, and also a face recognition subsystem which identifies the unknown face image. So the complete face recognition system as a system which interfaces with a video source, detects all face(s) images in each frame, and sends only the detected face images to the face recognition subsystem which in turn identifies the face images. The face detection subsystem uses our previously developed hardware implementation, which is publicly available at. The face recognition subsystem uses the Shape-from-shading algorithm.

3.2 Automatic Landmarking

Shape-from-shading (SFS) based method is proposed to generate synthetic facial images under different rotations and illuminations. Another study presents a combination of an edge model and color region model for face recognition, after synthetic images with varying pose are created via a deformable 3D model. Lately, a study in which 3D model reconstruction is achieved by applying the 3D Generic Elastic Model approach. Mage formulation rules tell you how to go from a 3D model and its materials to a 2D image. Shape from shading is the inverse problem. It can be seen as a constraint on the set of possible realities. Justifying working with it can take several arguments – the simplest is that multiple species of animals, ourselves included, use it Shape from Shading: First to work with entire surfaces, rather than 1D slices. A 3D generic face model is aligned onto a given frontal image using 115 feature vertices and different images are synthesized with variant pose, illumination and expression. An ultimate goal in computer vision is the 3-D reconstruction of our real world based on 2-D imagery. Although tremendous progress has been achieved. In this paper have address this so-called shape-from-shading (Sfs) problem by introducing a novel framework that is particularly tailored to the difficulties one has to face in real-world scenario method.

Disadvantages:

2D face recognition methods are sensitive to lighting, head orientations, facial expressions and makeup. 2D images contain limited information
3D Representation of face is less susceptible to isometric deformations (expression changes). 3D approach overcomes problem of large facial orientation changes

III. PROPOSED WORK

Address the problem of expressions in 2D probe images. Our aim is to facilitate recognition by simulating facial expressions on 3D models of each subject. With regard to the causes of the intra-class variations, synthesis of facial image under various pose and illumination conditions using 3D face models is straightforward since these variations are external.

In the proposed system, the enrolment 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 pre-processed, firstly to extract the face region. On the extracted facial surface, scanner-induced holes and spikes are cleaned and a bilateral 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 (4). These detected points are then utilized to warp a generic animatable face model so that it completely transforms into the target face.

The proposed work developed by using the following modules. Those are

1. Data Preprocessing
2. Automatic Landmarking
3. Animatable Face Models, Expression Simulations

3.1 Data Preprocessing

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

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

The existing spikes are removed by threshold. 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.

3.2 Automatic Landmarking

Bearing in mind that subject cooperation is required during the enrolment, 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.

On the other hand for the regions with noisy surface and/or distinctive colour 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). The steps are detailed in the following:

![FIG 5. AUTOMATIC LAND MARKING](image)

1) 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. 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.

2) 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 texture carries 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. For this purpose, the method proposed in our previous work is adopted. 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.

3) 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).

**Animatable Face Models**

In order to construct an animatable face model for each enrolled subject, a mesh warping algorithm based on the findings in is proposed. A generic face model, with holes for the eyes and an open mouth is strongly deformed to fit the facial models in the database, using the TPS method. 17 points to be automatically detected together with the rest of the FDP points for MPEG-4 compliant animations are annotated for the generic face. MPEG-4 specifications and the mathematical background of the TPS method will be briefly explained before going into details about the proposed animatable face construction method.
the positions of the MPEG-4 FDP points are given. Most of these points are necessary for an MPEG-4 compliant animation system, except for the ones on the tongue, the teeth and the ears, depending on the animation tool structure.

2) Thin Plate Spline Warping: As the name indicates, the TPS method is based on a physical analogy to how a thin sheet of metal bends under a force exerted on the constraint points. The TPS method was made popular by Fred L. Book stein in 1989 in the context of biomedical image analysis [3]. For the 3D surfaces S and T, and a set of corresponding points (point pairs) on each surface, Pi and Mi respectively, the TPS algorithm computes an interpolation function \( f(x,y) \) to compute \( T' \),

with \( U(.) \), the kernel function, expressed as:

\[
U(r) = r \ln r \quad r = x^2 + y^2 \quad (4)
\]

In the interpolation function \( f(x,y) \), the \( w_i \) s are the weights. As given in (3), the interpolation function consists of two distinct parts: An affine part \((a_1 + a_2x + a_3y)\) which accounts for the affine transformation necessary for the surface to match the constraint points and a warping part \((w_iU(|P_i - \{x,y\})|)\).

3) The Method: TPS is commonly used to establish registration in non-rigidly deformed surface patches, like two different facial surfaces. The deformation of the registered models is minimal since only few point pairs are utilized. As more control points are added to the TPS warping, the amount of deformation increases and the face becomes more and more similar to the target surface. By exploiting this fact, in this paper we propose to strongly deform a generic face to fit target faces in the gallery. On the other hand, before applying the TPS warping, we have to make sure that the target face and the generic face models are well-aligned. Each step of the process to obtain the animatable model is detailed in the following subsections.

IV. FACE RECOGNITION TECHNIQUES

Face Recognition performance is evaluated with three different techniques: Principal component analyses, Linear discriminate analysis and local binary patterns. Significant improvements are achieved in face recognition accuracies, for each database and algorithm.

1. Principal component analysis

Principal component analysis (PCA) is a technique that is useful for the compression and classification of data. The purpose is to reduce the dimensionality of a data set (sample) by finding a new set of variables, smaller than the original set of variables that nonetheless retains most of the sample's information. By information we mean the variation present in the sample, given by the correlations between the original variables. The new variables, called principal components (PCs), are uncorrelated, and are ordered by the fraction of the total information each retains.

2. local binary patterns

We empirically evaluate LBP features for person-independent facial expression recognition. Different machine learning methods are exploited to classify expressions on several databases. LBP features were previously used for facial expression classification in [4], and more recently, following our work [3], Liao et al. presented an extended LBP operator to extract features for facial expression recognition. However, these existing works were conducted on a very small database (JAFFE) using an individual classifier. In contrast, here we comprehensively study LBP features for facial expression recognition with different classifiers on much larger databases. We investigate LBP features for low-resolution facial expression recognition, a critical problem but seldom addressed in the existing work. We not only perform evaluation on different image resolutions, but also conduct experiments in real-world compressed video sequences. Compared to the previous work, LBP features provide just as good or better performance, so are very promising for real-world applications.

3. Linear discriminante analysis

Linear Discriminate Analysis, or simply LDA, is a well-known classification technique that has been used successfully in many statistical pattern recognition problems. It was developed by Ronald Fisher, who was a professor of statistics at University College London, and is sometimes called Fisher Discriminate Analysis (FDA). The primary purpose of LDA is to separate samples of distinct groups. We do this by transforming the data to a different space that is optimal for distinguishing between the classes. For example, let us suppose we have a simple two variable problem in which there are two classes of objects. Having made many measurements from examples of the two classes we can find the best Gaussian model for each class. The nominal class boundaries (where the probability of a point belonging to the class falls below a certain limit) are shown by the ellipses. Now, given a new measurement, we want to determination.

V. CONCLUSION

2D face recognition methods are sensitive to lighting, head orientations, facial expressions and makeup. 2D images contain limited information those type of problems overcome by using 3D image. Significant improvements are achieved in face recognition accuracies, for each database and algorithm. This model reveal that introduction of realistically synthesized face images with expressions improves the performance of the identification system.
VI. REFERENCES

[7] Exposing Digital Image Forgeries by IlluminationColor Classification Tiago José de Carvalho, Student Member, IEEE, Christian Riess, Associate Member, IEEE, Elli Angelopoulou, Member, IEEE, Hélio Pedrini, Member, IEEE, and Anderson de Rezende Rocha, Member, IEEE