www.ijecs.in
International Journal Of Engineering And Computer Science ISSN:2319-7242
Volume 3 Issue 1 January, 2014 Page No.3794-3804

RECOGNITION OF FACIAL EXPRESSION USING FACIAL MOVEMENT FEATURES

***D. Rama Krishna Reddy*[1], *D. Hema Latha*[2], *K. Sudha*[3]**

[1]Osmania University, Dept of Computer science,
College of Science, Hyderabad, India
drkreddymca@gmail.com

[2]Osmania University College for Women, Dept of Computer Science,
Koti, Hyderabad, India
hlatha059@gmail.com

3 sudhakatkuri@gmail.com

Abstract: This paper work presents an approach to solve recognition of facial expression using 'salient' distance features, which are obtained by extracting patch-based 2D Gabor features, selecting the 'salient' patches, and performing patch matching operation The experimental results demonstrate high correct recognition rate (CRR), significant performance improvements due to the consideration of facial +element and muscle movements, promising results under face registration errors, and fast processing time. The comparison with the state-of-the-art performance confirms that the proposed approach achieves the highest CRR on the JAFFE database and is among the top performers on the Cohn-Kanade (CK) database.

Keywords:

Introduction

Recognizing faces can be done very easily without much conscious thought, but still it has remained a difficult problem in the area of computer vision, where some 20 years of research is just beginning to yield useful technological solutions. As a biometric technology, automated face recognition has a number of desirable properties that are driving research into practical techniques.

The problem of face recognition can be stated as 'identifying an individual from images of the face' and encompasses a number of variations other than the most familiar application of mug shot identification. One notable aspect of face recognition is the broad interdisciplinary nature of the interest in it: within computer recognition and pattern recognition; biometrics and security; multimedia processing; psychology and neuroscience.

It is a field of research notable for the necessity and the richness of interaction between computer scientists and psychologists.

The automatic recognition of human faces spans a variety of different technologies. At a highest level, the technologies are best distinguished by the input medium that is used, whether visible light, infra-red [29, 31] or 3-dimensional data [7] from stereo or other range-finding technologies. Thus far, the field has concentrated on still, visible-light, photographic images, often black and white, though much interest is now beginning to be shown in the recognition of faces in color video. **Each input medium that is used for face recognition brings robustness to certain** conditions, e.g. infra-red face imaging is practically invariant to lighting conditions while 3-dimensional data in theory is invariant to head pose. Imaging in the visible light spectrum, however, will remain the preeminent domain for research and application of face recognition because of the vast quantity of legacy data and the ubiquity and cheapness of photographic capture equipment.

1.1 Face as a Biometric

Face recognition has a numerous strengths to recommend it over other biometric modalities in certain circumstances, and corresponding weaknesses that make it an inappropriate choice of biometric for other applications. Face recognition as a biometric derives a number of advantages from being the primary biometric that humans use to recognize one another. Some of the earliest identification tokens, i.e. portraits, use this biometric as an authentication pattern. Furthermore it is well-accepted and easily understood by people, and it is easy for a human operator to arbitrate machine decisions — in fact face images are often used as a human-verifiable backup to automated fingerprint recognition systems.

Because of its prevalence as an institutionalized and accepted guarantor of identity since the advent of photography, there are large legacy systems based on face images — such as police records, passports and driving licences — that are currently being automated. Video indexing is another example of legacy data for which face recognition, in conjunction with speaker identification [19], is a valuable tool.

Face recognition has the advantage of ubiquity and of being universal over other major biometrics, in that everyone has a face and everyone readily displays the face. (Whereas, for instance, fingerprints are captured with much more difficulty and a significant proportion of the population has fingerprints that can not be captured with quality sufficient for recognition.) Uniqueness, another desirable characteristic for a biometric, is hard to claim at current levels of Achievements and Challenges in Fingerprint Recognition accuracy. Since face shape, especially when young, is heavily influenced by genotype, identical twins are very hard to tell apart with this technology.

With some configuration and co-ordination of one or morecameras, it is be more or less possible to acquire face images without active participation of the subject. Such passive identification might be desirable for customization of user services and consumer devices, whether that be opening a house door as the owner walks up to it, or adjusting mirrors and car seats to the driver's presets when sitting down in their car. Surveillance systems rely on passive acquisition by capturing the face image without the cooperation or knowledge of the person being imaged. Face recognition also has the advantage that the acquisition devices are cheap and are becoming a commodity (though this is not true for non-visible wavelength devices and some of the more sophisticated face recognition technologies based on 3-dimensional data).

The main drawbacks to face recognition are its current relatively low accuracy (compared to the proven performance of fingerprint and iris recognition) and the relative ease with which many systems can be defeated (Section 4.2.1). Finally, there are many attributes leading to the variability of images of a single face that add to the complexity of the recognition problem if they can not be avoided by careful design of the capture situation. Inadequate constraint or handling of such variability inevitably leads to failures in recognition.

These include:

Physical changes: facial expression change; aging; personal appearance

(Make-up, glasses, facial hair, hairstyle, disguise).

Acquisition geometry changes: change in scale, location and in-plane rotation of the face (facing the camera) as well as rotation in depth (facing the camera obliquely, or presentation of a profile, not full-frontal face).

Imaging changes: lighting variation; camera variations; channel characteristics (especially in broadcast, or compressed images).

No current system can claim to handle all of these problems well. In particular there has been little research on making face recognition robust to the effects of aging the faces. In general, constraints on the application scenario and capture situation are used to limit the amount of invariance of face image sample that needs to be afforded algorithmically.

The main challenges of face recognition today are handling rotation in depth and broad lighting changes, together with personal appearance changes. Even under good conditions, however, accuracy needs to be improved.

1.2 Robustness and Fraud

All biometric recognition systems are susceptible to accidental errors of two types which both must be minimized: False Accept (FA) errors where a random impostor is accepted as a legitimate user and False Reject (FR) errors where a legitimate user is denied access. Designers of biometric systems must also be very conscious of how the system will behave when deliberately attacked.

Naturally much of biometric system design falls into the more traditional categories of physical, procedural and electronic security — preventing an attacker from circumventing the recognition system or preventing false enrolment of biometric identities into a system's database, for example. That is, purposeful and successful attempts at creating a false accept error by general means of security attacks. Nevertheless, there are a number of security attack types that are specific to biometrics.

It is very easy to change one's facial appearance to make one look very different, and so to prevent identification, i.e. cause a false rejection. This is particularly important in a 'non-cooperative' application where the biometric is being used to prevent a single person from obtaining a privilege (such as a vote or driving licence) more than once. While underlying bone structure is extremely difficult to change, it is also hard to measure, and all face recognition systems rely on more superficial, changeable characteristics making them defeasible for determined individuals.

It is also possible for some people to impersonate others with a high degree of

similarity (an important vulnerability in 'cooperative' applications like physical access control). Photographs, rubber masks, video replay all allow impostor attacks — the deliberate engineering of a false acceptance error. Detection of such fake biometrics data is only superficially handled by commercial systems, though this is improving. A couple of years ago, few systems had a test to detect authenticity (rejecting objects that looked too flat to be faces rather than photographs), but a recent PC Magazine test [21] found that both systems tested could distinguish a real person from a photograph. More sophisticated shape algorithms could be devised, and elastic deformation can be used to prevent simple photograph replay attacks. (One system allows the option of requiring a change in facial expression during verification.) With computing power more abundant, the technology for detecting fake biometrics will keep improving. Achievements and Challenges in Fingerprint Recognition The combination with other biometrics — particularly lip motion verification or speaker ID [23] reduces the exposure to impersonation attacks, but further measures are necessary to prevent video replay attacks where a pre-recorded sequence of the authorized individual is somehow injected into the system.

Well established in speaker identification literature [2], prompted-text or text independent verification can avoid a simple replay attack, at the cost of a more intrusive, complex and expensive system, but the advances in trainable speech and face synthesis algorithms furnish attacks on even these sophisticated systems.

1.3 The Technology of Face Recognition

In this section we briefly review some of the technologies that have been used for face recognition. In general, face recognition systems proceed by detecting the face in an image, with the effect of estimating and normalizing for translation, scale and in-plane rotation. Given a normalized image, the features, either global or local, are extracted and condensed in a compact face representation which can then be stored in a database or a smartcard and compared with face representations derived at later times.

1.3.1 Related Fields

Face recognition is closely related to many other domains, and shares a rich common literature with many of them. Primarily, face recognition relies upon face detection described in Section 4.3.2. For recognition of faces in video, face tracking is necessary, potentially in three dimensions with estimation of the head pose [18]. This naturally leads to estimation of the person's focus of attention [9, 32] and estimation of gaze [20] which are important in human computer interaction for understanding intention, particularly in conversational interfaces. Correspondingly there is much work on person tracking [27] and activity understanding [37] which are important guides for face tracking and for which face recognition is a valuable source of information. Recent studies have also begun to focus on facial expression analysis either to infer affective state [30] or for driving character animations particularly in MPEG-4 compression [26]. The recognition of visual speech (i.e. lip-reading, particularly for the enhancement of acoustic speech recognition) is also a burgeoning face image processing area [1].

1.3.2 Face Detection

Naturally, before recognizing a face, it must be located in the image. In some cooperative systems, face detection is obviated by constraining the user. Most systems use a combination of skin-tone and face texture to determine the location of a face and use an image pyramid to allow faces of varying sizes to be detected. Increasingly, systems are being developed to detect faces that are not full-frontal [13]. Cues such as movement and person detection can be used [38] to localize faces for recognition. Typically translation, scale and in-plane rotation for the face are estimated simultaneously, along with rotation-in-depth when this is considered.

1.3.3 Face Recognition

There is a great diversity in the way facial appearance is interpreted for recognition by an automatic system. Currently a number of different systems are under development, and which is most appropriate may depend on the application domain. A major difference in approaches is whether to represent the appearance of the face, or the geometry. Brunelli and Poggio [5] have compared these two approaches, but ultimately most systems today use a combination of both appearance and geometry. Geometry is difficult to measure with any accuracy, particularly from a single still image, but provides more robustness against disguises and aging. Appearance information is readily obtained from a face image, but is more subject to superficial variation, particularly from pose and expression changes. In practice for most purposes, even appearance-based systems must estimate some geometrical parameters in order to derive a 'shapefree' representation that is independent of expression and pose artefacts.

This is achieved by finding facial landmarks and warping the face to a canonical neutral pose and expression. Facial features are also important for geometric approaches and for anchoring local representations.

Face appearance representation schemes can be divided into local and global, depending on whether the face is represented as a whole, or as a series of small regions. Most global approaches are based on a principal components representation of the face image intensities. This representation scheme was devised first for face image compression purposes [17] and subsequently used for recognition purposes [39]. The latter coined the term eigenfaces for this type of representation. A face image is represented as a vector of intensities and this vector is then approximated as a sum of basis vectors (eigenfaces) computed by principal component analysis from a database of face images. These principal components represent the typical variations seen between faces and provide a concise encapsulation of the appearance of a sample face image, and a basis for its comparison with other face images. This principal components representation is, like for example the Fourier transform, a decorrelating transform to an alternative basis where good representations of the salient characteristics of an image can be created from only a few low-order coefficients despite discarding many of the higher-order terms. Achievements and Challenges in Fingerprint Recognition.

Other researchers have taken the approach of local representations. Local representations have the advantage that only part of the representation is corrupted by local changes on the face. Thus, donning sunglasses only affects the local features near the eyes, but it may still be possible to recognize someone from features derived from around the nose and mouth. However, as mentioned above, inherently local representations are harder to estimate and there is a trade-off between feature estimation precision and feature size (locality of the representation).

Matching. Having processed a face and extracted the features, these are stored or transmitted as a facial code (face template), which can be as small as 84 bytes (Visionics). For each representation type, a distance or similarity measure is defined that allows 'similar' faces to be determined. Much of the art in biometrics is in the design of a model of the biometric data and, given a scheme for extracting the model parameters as a representation of the data, in creating a similarity measure that correctly discriminates between samples from the same person and samples from different people. As with any biometric system, some threshold on similarity must be chosen above which two face images are deemed to be of the same person. Altering the threshold gives different False Accept and False Rejection Rates (Section 4.2.1) — trading the one off against the other depending on the security level required. This is a trade-off between convenience and security: user-friendly matchers have a low false reject rate, while secure matchers have a low false accept rate.

1.3.4 Performance

The Face Recognition Technology (FERET) tests from Jonathan Phillips provided an early benchmark of face recognition technologies. Phillips has continued the evaluation of face systems for US government agencies in the Face Recognition Vendor Tests [4]. This report provides an excellent independent evaluation of three state-of-the-art systems with concrete performance figures. The report highlights the limitations of current technology — while under ideal conditions performance is excellent, under conditions of changing illumination, expression, resolution, distance or aging, performance falls off, in some cases dramatically.

Current face recognition systems are not very robust yet against deviations from the ideal face image acquisition but there is continual performance improvement.

1.4 Privacy Issues

With the widespread deployment of security cameras, and the increasing financial and technological feasibility of automating this surveillance, fears have also increased about the potential for invasion of privacy that this technology can bring about. Notable deployments of face recognition in the London borough of Newham, in Tampa Florida and at the 2001 Super bowl have raised the spectre of intrusive applications of face recognition. It is now starting to become easy and cheap to connect a face recognition system to a blanket video surveillance system with great potential for crime prevention, but also bringing undreamt-of powers of control to totalitarian regimes, and the erosion of civil liberties by an ever-wakeful, omniscient 'big brother' capable of tracking the activities of its citizens from cradle to grave.

Technology will have answers to assuage these fears: Cryptography will go a long way toward privacy-guarding; and rigorous rights management, to limit access to the information, will prevent privacy violations by unauthorized individuals. Automatic identity-masking controls may make these technologies in theory less privacy-intrusive than human visual surveillance systems in that an automatic surveillance system can prevent voyeurism by only allowing people access to the video when a security incident has been detected. However, it seems that this technology is a tool as any other, and only legislation, self-regulation and social pressure will guide its use to beneficial rather than oppressive aims. Inevitably, in a pluralist world, there will be applications that tend to the latter.

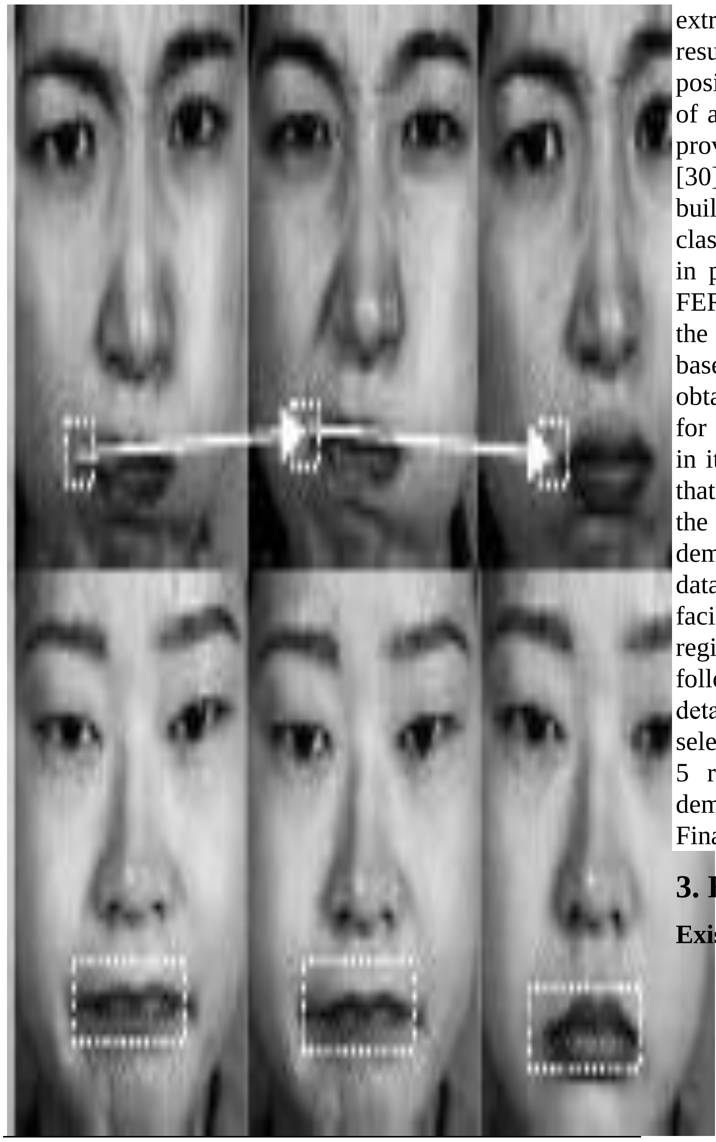
2. Literature Survey

The vast majority of the past work on FER does not take the dynamics of facial expressions into account [1]. Some efforts have been made on capturing and utilizing facial movement features, and almost all of them are video- based. These efforts try to adopt either geometric features of the tracked facial points (e.g. shape vectors [2], facial animation parameters [3], distance and angular [4], and trajectories [5]), or appearance difference between holistic facial regions in consequent frames (e.g. optical flow [6], and differential-AAM [7]), or texture and motion changes in local facial regions (e.g. surface deformation [8], motion units [9], spatiotemporal descriptors [10], animation units [11], and pixel difference [12]). Although achieved promising results, these approaches often require accurate location and tracking of facial points, which remains problematic [13]. In addition, it is still an open question how to learn the grammars in defining dynamic features, and handle ambiguities in the input data [14]. The image-based FER techniques provide an another approach to identify emotions based on appearance-based characteristics in a single image, and are essential for the situation where different images are available for training and testing.

The concept of patch matching operations are used to build characteristics or features for object recognition [15], [16] and action classification [17], which remain strong and efficient when there are changes in position, calibration or scale, and orientation. To apt for the aim of FER, the matching area and matching scale to restrict the operations within a suitable space are defined. By comparing and matching patch-based Gabor characteristics in this space, multi-distance values are obtained. The minimum distance is selected as the concluding feature for emotion classification. In this manner, one patch, which varies in its position, scale and shape, still can be captured provided that it is located within the defined matching space. To manifest the effectiveness and efficiency of using the proposed distance features, we describe the great and efficient functioning on two widely used databases, denoting enhancements based on the analysis of facial movement features, and favourable or promising results under face registration errors.

Facial expression recognition (FER) has been developed so much in recent years; the related fields in which development has occurred are especially machine learning, image processing and human cognition. Correspondingly, the impact and potential usage of automatic FER have been expanding in a wide range of applications, includes human-computer interaction, robot control and driver state surveillance. However, to date, robust recognition of facial expressions from images and videos is still a challenging work due to the trouble in exactly obtaining the useful emotional features. These characteristics are frequently represented in various forms, such as static, dynamic, point-based geometric or region-based appearance. *Facial movement features*, which include feature position and shape changes, are generally caused by the movements of facial elements and muscles during the course of emotional expression. The facial elements, especially key elements, will constantly change their positions when expressing emotions. As a sequel, the identical characteristic in different images has different positions, as shown in Fig.1 (a). In some circumstances, the shape of the feature may be deformed or garbled due to the deep facial movements of muscle. For example, the mouth in the first two images in Fig. 1 (b) represents different shapes from that in the third image. Therefore, for any feature representing a certain emotion, the geometric-based position and appearance-based shape normally changes from one image to other image in image databases, as well as in videos. This type of movement features represents an affluent pool of both static and dynamic characteristics of expressions that play a crucial role for FER.

The maximum of the previous work on FER does not take the active characteristics of facial expressions into account [1].but some facial movement features are captured and utilized, and almost all of them are video-based. These some captured features try to adopt either geometric features of the tracked facial points (e.g. shape vectors [2],facial animation parameters [3], distance and angular [4],and trajectories [5]), or appearance difference between holistic facial regions in consequent frames (e.g. optical flow [6], and differential-AAM [7]), or texture and motion changes in local facial regions (e.g. surface deformation[8], motion units [9], spatiotemporal descriptors [10],animation units [11], and pixel difference [12]). Despite of achieving favorable results, these methods require accurate location and tracking of facial points, which remains problematic [13]. In addition, it is still an open question how to learn the grammars in defining dynamic features, and handle ambiguities in the input data [14].Beside this, image-based FER techniques provide an alternative way to recognize emotions based on appearance-based features in a single image, and are important for the situation where only several images are



1) *Figure - 1*

Fig. 1. Facial movement features. (a) Feature position (left mouth corner) changes. (b) Feature shape (mouth) changes. Facial regions are manually cropped from two subjects “KA” and “KL” on the JAFFEdatabase.

available for training and testing. However, to the best of our knowledge, no research has been reported on image-based FER that considers facial movement features. In this paper, we aim for improving the performance of FER by automatically capturing facial movement features in static images based on distance features. The distances are obtained by extracting ‘salient’ patch-based Gabor features and then performing patch matching operations. Patch-based Gabor features have shown excel-lent performance in overcoming position, scale, and orientation changes [15], [16], [17], as well as extracting spatial, frequency and orientation information [18]. They also show a great advantage over the commonly used fiducially point-based Gabor [19], [20], [21], [22], [23],graph-based Gabor [24] and discrete Fourier transform[25] features in capturing regional information. Although other appearance-based features, such as local binary pat-terns (LBP) [26], [27], [10], Haar [28] and histograms of oriented gradients (HOG) [29], have shown a good performance in FER, they lack the capacity of capturing facial movement features with high accuracy. This is due to the fact that these appearance-based features are based on statistic values (e.g. histogram similarity) extracted from sub-regions; therefore, they produce similar results even when facial features move a bit from the original position. On the other hand, Gabor features have the capacity of accurately capturing movement information, and have been proven as being robust even in the case of face misalignment [30].The idea of patch matching operations has been used to build features for object recognition [15], [16] and ac-tion classification [17], which remain robust when there are changes in position, scale, and orientation. To fit for the purpose of FER, we define the matching area and matching scale to restrict the operations within a suitable space. By matching patch-based Gabor features in this space, multi-distance values are obtained. The minimum distance is chosen as the final feature for emotion classification. In this way, one patch, which varies in its position, scale and shape, still can be captured provided that it is located within the defined matching space. To show the effectiveness of using the proposed distance features, we demonstrate the high performance on two widely used databases, significant improvements due to the consideration of facial movement features, and promising results under face registration errors. The remainder of the paper is organized as follows. Section 2 describes the proposed framework, while the details of building distance features and ‘salient’ feature selection are explained in Section 3 and 4 respectively. Section 5 represents the recognition and speed performance, and demonstrates comparison with the state-of-the-art performance. Finally, conclusions are drawn in Section 6.

3. Existing System

Existing System

- The vast majority of the past work on FER does not take the dynamics of facial expressions into account.
- Some efforts have been made on capturing and utilizing facial movement features, and almost all of them are video-based.
- These efforts try to adopt either geometric features of the tracked facial points (e.g. shape vectors, facial animation parameters, distance and angular, and trajectories, or appearance difference between holistic facial regions in consequent frames (e.g. optical flow, and differential-AAM, or texture and motion changes in local facial regions (e.g. surface deformation, motion units, spatiotemporal descriptors, animation units, and pixel difference).
- Although achieved promising results, these approaches often require accurate location and tracking of facial points, which remains problematic.

4. Proposed Work

The proposed work is composed of **pre-processing, training and test stages**. In the pre-processing stage, the nose is taken as the centre point and keeping essential facial components inclusive, the facial regions are manually gathered from the database images and calibrated or scaled to a resolution of 48*48 pixels. To imitate the results of real face detectors no other processing is conducted.

Then multi-resolution Gabor images are achieved by convolving eight-scale, four-orientation Gabor filters with the calibrated or scaled facial regions. During the training stage, a entire set of patches are derived or taken from by moving a series of patches with different sizes across the training Gabor image

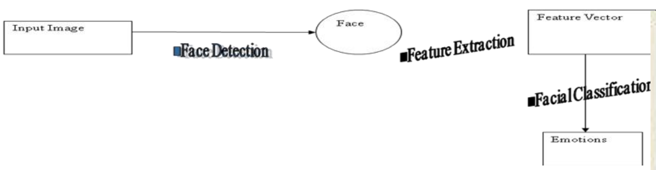
Then patch matching operation is propounded to transform the derived patches to distance characteristics. To capture facial movement characteristics, the matching domain and calibration are defined to enhance the matching space, whereas the minimum criteria is used to discover the best matching feature in this space. Based on the transformed distance characteristics, a set of 'salient' patches is chosen by Adaboost.

In the test stage, the same patch matching operation is carried out on a new image using the 'salient' patches. The resulting distance characteristics are given as input to a multi-class support vector machine (SVM) to identify six basic emotions, including anger (AN), disgust (DI), fear (FE), happiness (HA), sadness (SA) and surprise (SU).

The main aim of the technique we used is that it takes static images of a single user as input, and finds the expression made with a high rate of accuracy. The static images are taken as input from a webcam and are processed to satisfy the functional requirements as given below.

1. **Face localization**: To detect the existing face that is taken as an image from the webcam.
2. **Region of Interest Detection**: To partition the face into regions of interest for feature extraction in each region.
3. **Feature Extraction**: To detect the characteristics present in each region of interest.
4. . **Feature Classification**: To classify the characteristics detected based on their relative and absolute distances on the face.5.
5. **pression Recognition**: To identify the expression made based on the feature movements

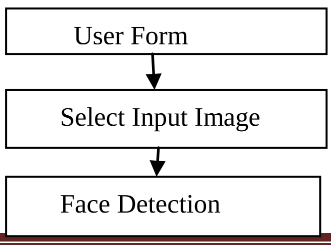
6. System Flow Diagram

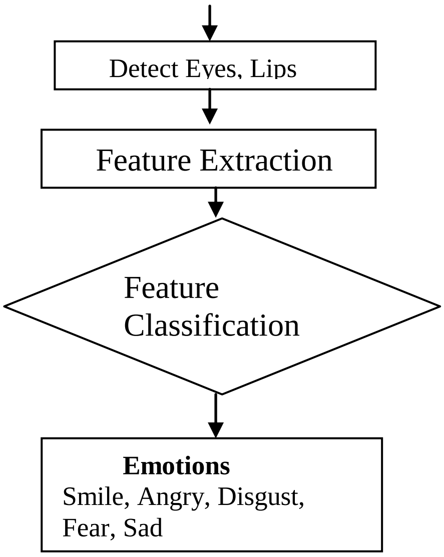


7.

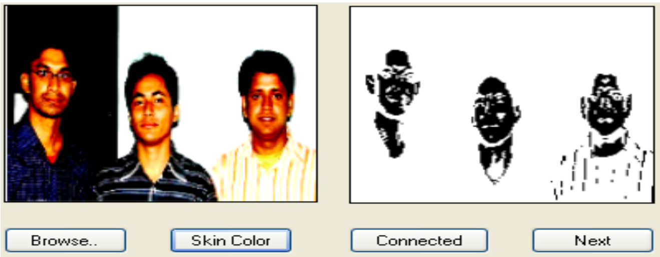
The System Flow Diagram explains the total flow in terms of phases. Three main phases: Actual Face detection, the Features (eyes, lips) Extraction, Facial Classification (Emotions).

5.1 Data Flow Diagram





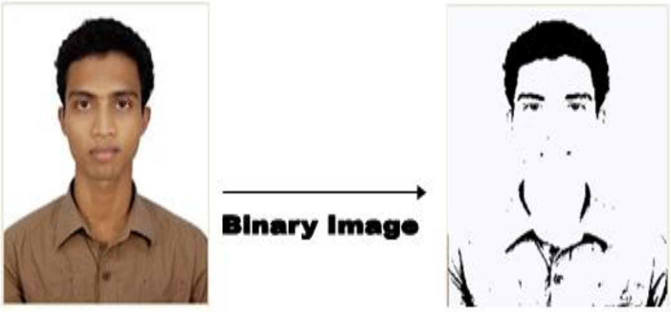
For skin color segmentation, first the image is compared. Then skin color segmentation is performed.



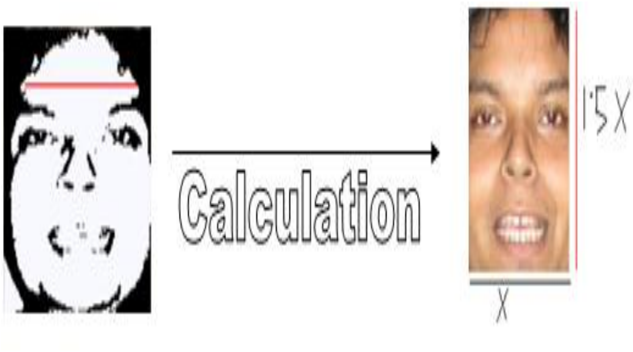
Then, find the largest connected region. And next check the probability to develop a face of the largest connected region. If the largest connected region has the probability to develop a face, then it will open a new form with the largest connected region. If the largest connected regions elevation and breadth is larger or equal than 50 and the ratio of elevation/breadth is between 1 to 2, then it may be face.

Face Detection

For face detection, first transform the binary image from RGB image. For transforming binary image, calculate the average value of RGB for each pixel and if the average value is below than 110, we substitute it by black pixel else substitute it by white pixel. By this method, binary image is extracted from RGB image.



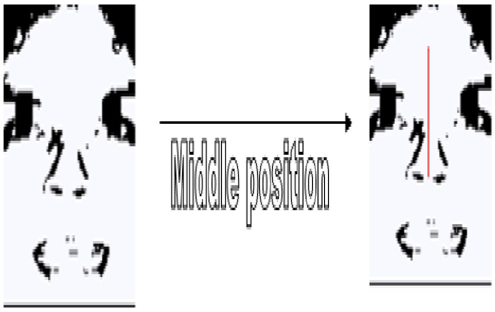
Then, next step is to discover the forehead from the binary image. Now image is scanned. Scanning is done from the middle of the image, then need to find continued sequence of white pixels after the continued sequence of black pixels. Then we need to discover the maximum breadth or width of the white pixel by searching vertical both left and right site. Then, if the new width is smaller than the half of the previous maximum width, then break the scan because if we reach the eyebrow then this situation will arise. Then cut the face from the starting position of the forehead and its height will be multiplied 1.5 times of its width.



In the figure, X will be equal to the maximum width of the forehead. Then we get an image which contains only eyes, nose and lip. Then the RGB image is cut according to the binary image.

Eyes Detection

For detecting eyes, the RGB face is converted into binary face. And now, consider the face width by W and scan the face from the W/4 to (W-W/4) to locate the centre position of the two eyes. The highest white continuous pixel along the height between the ranges is the middle position of the two eyes.



Then discover the starting upper position of the two eyebrows by searching vertically. For left eye, search w/8 to mid and for right eye search mid to w – w/8. Here w - width of the image, mid - middle position of the two eyes.

There may be few white pixels between the eyebrow and the eye. Then place some continuous black pixels vertically from eyebrow to the eye, to have connectivity between the eyebrow and the eye. For left eye, the vertical black pixel-lines are placed in between mid/2 to mid/4 and for right eye the pixel lines are in between mid+(w-mid)/ 4 to mid+3*(w-mid)/ 4 and height of the black pixel-lines are from the eyebrow starting height to (h- eyebrow starting position)/4. Here w is the width of the image and mid is the middle position of the two eyes and h is the height of the image. Then we find the lower position of the two eyes by searching black pixel vertically.

For left eye, we search from the mid/4 to mid - mid/4 width. And for right eye, we search mid + (w-mid)/ 4 to mid+3*(w-mid)/ 4 width from image lower end to starting position of the eyebrow. Then discover the right side of the left eye by searching black pixel horizontally from the mid position to the starting position of black pixels in between the upper position and lower position of the left eye. And left side for right eye we search mid to the starting position of black pixels in between the upper position and lower position of right eye. The left side of the left eye is the starting width of the image and the right side of the right eye is the ending width of the image. Then we cut the upper position, lower position, left side and the right side of the two eyes from the RGB image.

Lip Detection

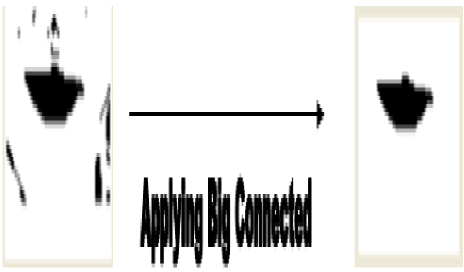
For lip detection, check the lip box. And we consider that lip must be interior to the lip box. So, first arbitrate or determine the distance between the forehead and eyes. Then sum up the distance with the lower height of the eye to determine the upper height of the box that contains the lip. Now, the initial point of the box will be the ¼ position of the left eye box and end point will be the ¾ position of the right eye box. And the ending height of the box will be the lower end of the face image. So, this box will contain only lip and some part of the nose. Then RGB image is cut according the box.

So, for eyes and lips detection, we only need to transform RGB image to binary image and some search among the binary image.

Apply Bezier Curve on Lip

In the lip box, lip and some part of nose is available. So, around the box there is skin color or the skin. So, transform the skin pixels to white pixels and other pixels as black pixels. Also identify the pixels that are similar to skin pixels and transform to white pixels. Here, if the difference of two pixels RGB value is less than or equal 10, then we call them similar pixel. Here, histogram can be used to discover the distance between the lower average RGB value and higher average RGB value. If the distance is less than 70, then we use 7 for finding similar pixel and if the distance is getter than or equal 70 then we use 10 for finding similar pixel. So, the value for finding similar pixel depends on the quality of the image. If the image quality is high, we use 7 for finding similar pixel and if the image quality is low, we use 10.

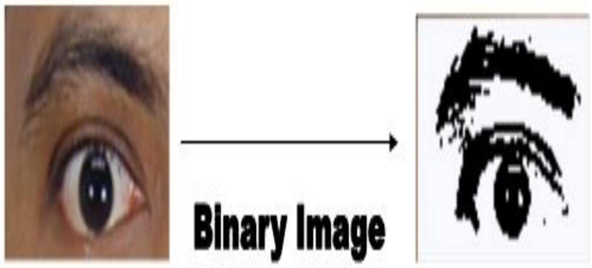
So, in the binary image, there are black regions on lip, nose and may some other little part which have a little different than skin color. Then we apply big connected region for finding the black region which contain lip in binary image. And we are sure that the big connected region is the lip because in the lip box, lip is the largest thing which is different than skin.



Then we have to implement or assign Bezier curve on the binary lip. To implement Bezier curve, we should discover the beginning and ending pixel of the lip in horizontal. Then two tangents are drawn on upper lip from the beginning and ending pixel and also discover two points on the tangent which is not the portion or component of the lip. For the bottom lip, we find two point analogous or similar process of the top lip. We use Cubic Bezier curves to depict or design the Bezier curve of the lip. We depict or design two Bezier curve for the lip, one for top or above lip and one for bottom or below lip.

Implement Bezier Curve on Eye

To implement or assign Bezier curve on eyes, first the eyebrow must be removed or deleted from eye. To delete eyebrow, we explore or check 1st continued black pixel then continued white pixel and then continued black pixel from the binary image of the eye box. Then the 1st continuous black pixel from the box is deleted and then we get the box which consists only the eye.



Now, the eye box which consist only eye has some skin or skin color around the box. So, we implement or assign similar skin color as implemented for the lip to discover the region of eye. Then we assign or apply big link to discover the highest connected region and this is the eye because in the eye box, eye is the biggest thing which is not similar to the skin color.

Then we assign the Bezier curve on the eye box, as done for to the lip. Then we get the mould or shape of the eye.

Database and Training

In this database, two tables are maintained. One table "`Person`" is for storing the name of persons and the other table "Position" which for storing index that consist of four kinds of emotions of persons. In the "`Position`" table, for each index, there are six control or command points for lip Bezier curve, 6 control or command points for left eye Bezier curve, 6 control or command points for right eye Bezier curve, lip elevation and breadth, left eye height and width and right eye height and width. So, by this methodology, the emotions of the people are extracted.

Emotion Detection

To detect emotion in an image, the Bezier curve of the lip, left eye and right eye is discovered. Then convert each width of the Bezier curve to 100 and elevation according to its breadth. If the person's emotion information is available in the database, then the algorithm will match which emotion's elevation is nearest the current elevation and the algorithm will give the nearest emotion as output.

If the person's emotion information is not obtainable in the database, then the algorithm calculates the average elevation or height for each emotion in the database for all persons and then come to a decision according to the average elevation.

6. CONCLUSION AND FUTURE WORK

This paper addresses the issue of facial expression recognition using facial movement features. The efficiency and capability of the propound approach is asserted by the recognition performance, computational time, and comparison with the latest or up-to-date performance. The empirical results also exhibit significant performance enhancements due to the consideration of facial movement features, and assuring performance under face registration errors. The outcome signify that patch-based Gabor characteristics or attributes reveals a higher quality performance when compared with point-based Gabor features in terms of deriving local characteristics, keeping the position information, acquiring a high quality recognition performance, and desiring for a low number. Various emotions have various'salient' areas; however, the most of these domains are spread and concentrated around mouth and eyes. In addition, these 'salient' domains or regions for each emotion seem to be not influenced by the choice of using point-based or using patch-based features. The 'salient' patches are distributed across all scales with an emphasis on the higher scales. For both the JAFFE and CK databases, DL2 performs the best among four distances. As far as emotions are considered, anger contributes or bestows most to the misrecognition. The JAFFE database needs greater sizes of patches than the CK database to store useful information. The proposed approach can be potentially applied for various applications, such as detecting patient medical condition, monitoring driver or exhaust level or fatigue, and intelligent instructing system.

Future Work In future, the work can be extended to a video-based FER system by combining patch-based Gabor features with motion information in multi-frames. Recent progress on action recognition [47] and face recognition [48] has laid a foundation for using both appearance and motion features.

7.References:

[1] Z. Zeng, M. Pantic,G. I. Roisman, and T. S. Huang, A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions, *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol. 31, pp. 39 58, 2009.
[2] T. Yan, C. Jixu, and J. Qiang, A Unified Probabilistic Framework for Spontaneous Facial Action Modeling and Understanding, *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol. 32, pp. 258 273, 2010.
[3] P. S. Aleksic and A. K. Katsaggelos, Automatic facial expression recognition using facial animation parameters and multi stream HMMs, *Information Forensics and Security, IEEE Transactions on,* vol. 1, pp. 3 11, 2006.
2) *[4] D. Hamdi, V. Roberto, S. Albert Ali, and G. Theo, Eyes do not lie: spontaneous versus posed smiles, in Proceedings of the international conference on Multimedia Firenze, Italy: ACM, 2010, pp.703 706.*
[5] T. H. Wang and J. J. James Lien, Facial expression recognition system based on rigid and non rigid motion separation and 3D pose estimation, *Pattern Recognition,* vol. 42, pp. 962 977, 2009.
[6] M. Yeasin, B. Bullot, and R. Sharma, Recognition of facial expressions and measurement of levels of interest from video, *Multimedia, IEEE Transactions on,* vol. 8, pp. 500 508, 2006.
[7] Y. Cheon and D. Kim, Natural facial expression recognition using differential AAM and manifold learning, *Pattern Recognition,* vol. 42, pp. 1340 1350, 2009.

[8] F. Tsalakanidou and S. Malassiotis, Real time 2D+3D facial action and expression recognition, *Pattern Recognition,* vol. 43, pp. 1763 1775, 2010.
[9] I. Cohen, N. Sebe, A. Garg, L. S. Chen, and T. S. Huang, Facial expression recognition from video sequences: temporal and static modeling, *Computer Vision and Image Understanding,* vol. 91, pp. 160 187, 2003.
[10] G. Zhao and M. Pietikainen, Boosted multi resolution spatio temporal descriptors for facial expression recognition, *Pattern Recognition Letters,* vol. 30, pp. 1117 1127, 2009.
[11] F. Dornaika and F. Davoine, Simultaneous Facial Action Track
ing and Expression Recognition in the Presence of Head Motion, *International Journal of Computer Vision,* vol. 76, pp. 257 281, 2008.
[12] A. Kapoor, W. Burleson, and R. W. Picard, Automatic prediction of frustration, *International Journal of Human Computer Studies,* vol. 65, pp. 724 736, 2007.
[13] L. Peng and S. J. D.Prince, Joint and implicit registration for face recognition, in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on,* 2009, pp. 1510 1517.
[14] T. Huang, A. Nijholt, M. Pantic, and A. Pentland, Human Computing and Machine Understanding of Human Behavior: A Survey, in *Artifical Intelligence for Human Computing*. vol. 4451: Springer Berlin / Heidelberg, 2007, pp. 47 71.
[15] T. Serre, L. Wolf, S. Bileschi, M. Riesenhuber, and T. Poggio, Robust Object Recognition with Cortex Like Mechanisms, *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol. 29, pp. 411 426, 2007.
[16] J. Mutch and D. G. Lowe, Multiclass Object Recognition with Sparse, Localized Features, in *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on,* 2006, pp. 11
[17] H. Jhuang, T. Serre, L. Wolf, and T. Poggio, A Biologically Inspired System for Action Recognition, in *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on,* 2007, pp. 1 8.
[18] L. Zhen, L. Shengcai, H. Ran, M. Pietikainen, and S. Z. Li, Gabor volume based local binary pattern for face representation and recognition, in *Automatic Face & Gesture Recognition, 2008. FG 08. 8th IEEE International Conference on,* 2008, pp. 1 6.
[19] L. Wiskott, J. M. Fellous, N. Kuiger, and C. von der Malsburg, Face recognition by elastic bunch graph matching, *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol. 19, pp. 775 779, 1997.
[20] G. Littlewort, M. S. Bartlett, I. Fasel, J. Susskind, and J. Movellan, Dynamics of facial expression extracted automatically from video, *Image and Vision Computing,* vol. 24, pp. 615 625, 2006.
[21] H. Y. Chen, C. L. Huang, and C. M. Fu, Hybrid boost learning for multi pose face detection and facial expression recognition, *Pattern Recognition,* vol. 41, pp. 1173 1185, 2008.
[22] S. Hoch, F. Althoff, G. McGlaun, and G. Rigoll, Bimodal fusion of emotional data in an automotive environment, in *Acoustics ,Speech, and Signal Processing, 2005. Proceedings. (ICASSP 05). IEEE International Conference on,* 2005, pp. 1085 1088.
[23] G. Guo and C. R. Dyer, Learning from examples in the small sample case: face expression recognition, *IEEE Transactions on Systems, Man, and Cybernetics, Part B,* vol. 35, pp. 477 488, 2005.
[24] S. Zafeiriou and I. Pitas, Discriminant Graph Structures for Facial Expression Recognition, *Multimedia, IEEE Transactions on,* vol. 10, pp. 1528 1540, 2008.
[25] T. Xiang, M. K. H. Leung, and S. Y. Cho, Expression recognition using fuzzy spatio temporal modeling, *Pattern Recognition,* vol. 41, pp. 204 216, 2008.
[26] C. Shan, S. Gong, and P. W. McOwan, Facial expression recognition based on Local Binary Patterns: A comprehensive study, *Image and Vision Computing,* vol. 27, pp. 803 816, 2009.
[27] Z. Guoying and M. Pietikainen, Dynamic Texture Recognition Using Local Binary Patterns with an Application to Facial Expressions, *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol. 29, pp. 915 928, 2007.
[28] P. Yang, Q. Liu, and D. N. Metaxas, Boosting encoded dynamic features for facial expression recognition, *Pattern Recognition Letters,* vol. 30, pp. 132 139, 2009.
[29] C. Orrite, A. Gañán, and G. Rogez, HOG Based Decision Tree for Facial Expression Classification, in *Pattern Recognition and Image Analysis,* 2009, pp. 176 183.
[30] S. Shiguang, G. Wen, C. Yizheng, C. Bo, and Y. Pang, Review the strength of Gabor features for face recognition from the angle of its robustness to mis alignment, in *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on,* 2004, pp. 338 341 Vol.1.
[31] D. Gabor, Theory of communication, *Institution of Electrical Engineers Journal Radio and Communication Engineering,* vol. 93, pp. 429 457, 1946.
[32] C. Cortes and V. Vapnik, Support vector networks, *Machine Learning,* vol. 20, pp. 273 297, 1995.
[33] C. C. Chang and C. J. Lin, LIBSVM: a library for support vector machines, 2001, *Software available at http://www.csie.ntu.edu.tw/cjlin/libsvm,* 2001.
[34] T. Serre, L. Wolf, and T. Poggio, Object recognition with features inspired by visual cortex, in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on,* 2005, pp. 994 1000 vol. 2.
[35] M. Lyons, S. Akamatsu, M. Kamachi, and J. Gyoba, Coding facial expressions with Gabor wavelets, in *Automatic Face and Gesture Recognition, 1998. Proceedings. Third IEEE International Conference on,* 1998, pp. 200 205.
[36] T. Kanade, J. F. Cohn, and T. Yingli, Comprehensive database for facial expression analysis, in *Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on,* 2000, pp. 46 53.
[37] Y. Freund and R. E. Schapire, A decision theoretic generalization of on line learning and an application to boosting, 1997, pp. 119 139.
[38] M. S. Bartlett, G. Littlewort, I. Fasel, and J. R. Movellan, Real Time Face Detection and Facial Expression Recognition: Development and Applications to Human Computer Interaction, in *Computer Vision and Pattern Recognition Workshop,*