RESEARCH ARTICLE OPEN ACCESS

Facial Expressions Recognition Based On Modified Action Units Classification

Felix Kipchumba Chepsiror [1], P. Shanmugavadivu [2]
Department Of Computer Science and Applications
Gandhigram Rural Institute-Deemed University
Gandhigram-624302
Tamil Nadu – India

ABSTRACT
Face recognition has become one of the most researched field over the past few decades. It has become very essential especially in surveillance, criminal identification and biometric identification. This paper does a review on widely applied feature oriented techniqniques for face recognition.
Keywords:- Face recognition, feature extraction, methods

I. INTRODUCTION

Research in the field of face recognition has been carried out for many years. Initial research done on face recognition dates back to 1950s. Moreover, other applications of facial recognition include surveillance, e-learning and robotic human-machine interfaces. In this paper, various methods that are used in facial recognition are explored. These methods are: Geometric feature based methods, knowledge based technique, template matching, feature invariant technique, appearance based methods ,global feature extraction techniques, spatio-temporal information based approaches, statistic model based approaches and feature vector based methods.

II. METHODS

The methods largely used in face recognition classification are:

Geometric Feature Based Methods

This technique is used to extract features such as eyes, mouth and nose. A lot of researches have been done using this approach, Ekman and Friesen [3] proposed the FACS system that describes movements on the face, where 44 Action Units(AU) are defined, and each one represents a movement of a particular part of the face. According to Ekman and Friesen [3], a facial expression is characterized by a combination of AUs.Hammal et al [4] developed a classifying system based on the belief theory, and applied it on the Hammal Caplier database [5]. They used five distances between three essential parts of the face (two eyes and mouth). In their work, distances are computed on skeletons of expression but only four emotions (joy, surprise, disgust and neutral) are considered from the six basic emotions. Abdat et al [6] used twenty one distances between all parts of the face to encode a facial expression .They use the variation of muscle relative to the neutral state, and for the classification method they used a statistical classifier, named Support Vector Machine (SVM). This method was tested on images from the Cohn-Kanade database [7] and FEEDTUM database [8]. In their work, an important number of parameters is used in the whole of the face which is laborious and time consuming.

Table (1): Comparison of Geometric based methods

Method	Advantage	Disadvantage
Hammal et al	Faster computation	Two eyes and mouth only considered. Four emotions classified:surprise,joy,disgust,neutral
Abdat et al	All the six basic emotions classified	Laborious and time consuming

III. KNOWLEDGE BASED TECHNIQUE

This method is based on the relationship between different facial features and their locations. These features are eyes, nose and mouth. However, this method has pros and cons as summarised in the table below. The advantage is that all facial features are individually extracted, while its disadvantage is that some facial features cannot be extracted due to changes in facial orientation.

Table (2): Advantage and Disadvantage of knowledge based technique

Advantage	Disadvantage
Individual facial features extracted	Affected by changes in facial orientation

IV. FEATURE INVARIANT TECHNIQUE

et al[9],compared different template matching techniques for face recognition.

Table (3): Comparison of Template Matching Methods

Template matching method	Accuracy(%)	Clutter Background(%)
Optimized Sum of Absolute Difference(OSAD)	100	96
Optimized Sum Squared of Difference(OSSD	98	92
Sum of Absolute Difference(SAD)	98	94
Sum of Squared difference(SSD)	95	89
Normalized Cross Correlation(NCC)	80	73
Zero Normalized Cross	80	73

ADANTAGES OF PCA

1. Receognition is simple and effective.

2. Data compression is attained by the little dimensional subspace depiction

Feature invariant approach makes it possible to recognize existing structural features even when pose, view or lighting conditions change. To cope with the facial changes which often appear only on some regions of the whole image due to variations in facial expression, illumination condition, pose, etc., a face image is divided into a number of non-overlapping sub-images. Each of these sub-images tries to represent local facial changes. The problem with this method is occlusion: this means the feature you are focusing on is obscured or hidden.However,the benefit of using feature based approach includes high speed matching.

V. TEMPLATE MATCHING

In this method the principle is to calculate the correlation or matching between areas of the input image and face previously created. Several patterns are stored to describe the face as a whole or the facial features separately The disadvantage of this method is scale variation.

There are several template matching techniques that have been proposed. Nadi

Correlation(ZNCC)		
Sum of Humming Distance(SHD)	43	40

From this comparison, they found out that Optimized Sum of Absolute Difference has 100% accuracy, therefore it is the best method for template matching.

APPEARANCE BASED METHODS

In this method, a supervised learning technique is used to determine whether an image belongs to the class of faces or non faces. However the problem with this method is the training samples. In [10] an appearance based local approach for feature extraction to overcome the setbacks of Principal Component Analysis was proposed. They converted colour images into gray scale images to overcome the problem of color.The Principal Component Analysis is a flexible reduction process. It is valuable when an individual has gained a large number of variables

3. Raw intensity statistics are used openly for learning and recognition without any major low-level or mid-level processing.

4. No information of geometry and reflectance of faces is compulsory

DISADVANDAGES OF PCA

1. The technique is very profound to scale, therefore, a low-level pre-processing is essential.

2. Its recognition rate falls for recognition beneath changing posture and lighting.

3. The problem can be more challenging when, great change in posture as well as in appearance occurred.

4. Learning is very slow, which makes it tough to modernize the face dataset.

VI. GLOBAL FEATURES EXTRACTION

In this technique, the whole face is considered for extraction. It covers all features of the face such as mouth, eyes and nose. Neeta et al[12] used five spatio temporal features for each image and the features are distance of eyebrows (vd0),distance between right eyebrow and nose tip(vd1),distance between left eyebrow and nose tip(vd2),mouth width(vw),mouth, height(vh).These features are used to create feature vector for classification. Tasnim et al[10] used six features for feature extraction. These are :

1 Vd=distance between right eye and nose tip

2. Ve=distance between left eye and nose tip.

3.Vh=mouth height.

4.Vw=mouth width

5.Vnm=distance between mouth and nose tip

6.Ve=distance between eyebrows.

Values from the detected parts are measured using Euclidian formula`.$ED=\sqrt{((X1-X2)^2+(Y1+Y2)^2)}$ (1)

Where, (X1, Y1) is the detected point of a facial part and (X2, Y2) is another detected part of facial part as well as ED is the distance between those detected facial parts. After calculating those features, calculated the mean of each feature for trained images. Then Canberra Distance (CD) is used as classifier. Where,

$$CD = jX1 - X2j \quad (2)$$
$$jX1j + jX2j$$

If the distance between two features is minimum then they seem to be similar. Here, X1 and X2 indicates two features. Jeemoni *et al.* [11] use Euclidean distance based decision-making technique to get the minimum distance of each features of trained images and testing images. On the other hand, Tasnim et al[10]have used CD in spite of ED because ED return sonly the distance between two points but CD measures the distance along with the vector of two points. That ensures the great measurement of facial expression recognition.

VII. APPROACHES OF VIDEO-BASED FACE RECOGNITION

SPATIO -TEMPORAL INFORMATION BASED APPROACHES

There are several algorithms that are used to extract 2D and 3D videos. The distance between two videos is the minimum distance between two frames across two videos. Zhou and Chellappa presented a sequential importance sampling (SIS) method to incorporate temporal information in a video sequence for face recognition [1], it nevertheless considered only identity consistency in temporal domain and thus it may not work well when the target is partially occluded. In [2], Krueger and Zhou selected face sample images as from training videos by on-line version of radial basis functions. This model is effective in capturing small 2D motion but it may not deal well with large 3D pose variation or occlusion. The condensation algorithm could be used as an alternative to model the temporal structures [3].

ADVANTAGE

- It is effective in capturing 2D motion.

DRAWBACKS

- Local information is not well exploited
- Intrapersonal information which is related to facial expression and emotions is encoded and used.
- Equal weights are given to the spatio-temporal features despite the fat that some of the features are more than others.
- A lot of methods can only handle well aligned faces thus limiting their use in practical scene.
- SIS method may not work well due to occlusion.

VIII. STATISTICAL MODEL BASED APPROACHES

In [1],models from videos were obtained by using low level feature techniques such as principal component analysis from images, which was used for matching a single frame and a video stream or between two video streams. Principal component null space analysis (PCNSA) is proposed in [4], which is helpful for non-white noise covariance matrices. Recently, the Autoregressive and Moving Average (ARMA) model method is proposed in [5] to model a moving face as a linear dynamical object. S. Soatto, G. Doretto, and Y. Wu proposed dynamic textures for video-based face recognition. HMM has been applied to solve the visual constraints problem for face tracking and recognition [6].

ADVANTAGES

- Principal Component Null Space Analysis (PCNSA) is helpful for non-white noise covariance matrices.
- HMM solves the visual constraints problem for face tracking and recognition.

IX. HYBRID CUES

There are methods that utilize other cues obtained from a video such as voice, mouth and gait. In [7] two cues: face and gait were combined which resulted in increased performance. [8] used face and speaker recognition techniques for audio-video biometric recognition. The paper combined histogram normalization, boosting technique and a linear discrimination analysis to solve problems such as illumination, pose and occlusion and proposes an optimization of a speech denoising algorithm on the basis of Extended Kalman Filter(EKF). In [9], Radial basis function neural networks approach uses face and mouth features to recognize a person in video sequences.

ADVANAGES

- High performance
- A combination of histogram normalization, boosting technique and linear discriminant analysis solves the problem of illumination, pose and occlusion.

X. FEATURE VECTOR BASED METHODS

In these methods, feature vectors are extracted from input videos, which are used to match with all the videos in the database.

1. **KNOWLEDGE BASED(TOP-DOWN) APPROACH**
   The relationship between facial features is captured to represent the contents of a face and encode it as a set of rules.
2. **FEATURE INVARIANT(BOTTOM-UP) APPROACH**
   Features such as face, mouth, nose and eyes are considered in this approach. Color-based approach makes use of the fact that the skin color can be used as indication to the existence of human using the fact that different skins from different races are clustered in a single region.
3. **FACIAL FEATURES BASED APPROACH**
   In this approach facial features are examined to find out whether an image belongs to a human face. The face texture is tested by using Space Gray Level Dependency (SGLD) matrix.

ADVANTAGE OF FEATURE VECTOR BASED METHODS

- All facial features are extracted

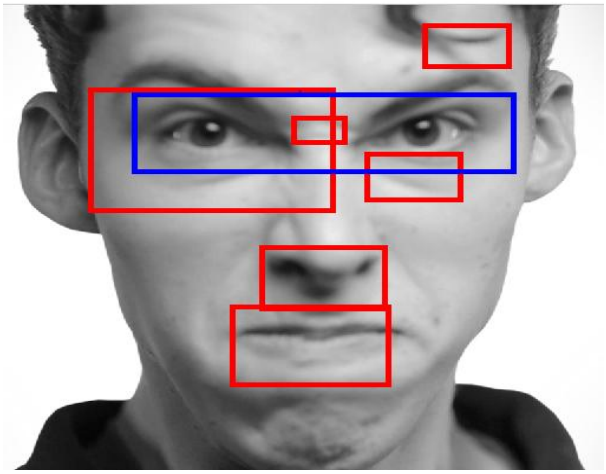
DISADVANTAGE

- Spatial information of input videos is neglected, which limits the performance of feature vector based approaches.

XI. MODIFIED ACTION UNITS CLASSIFICATION

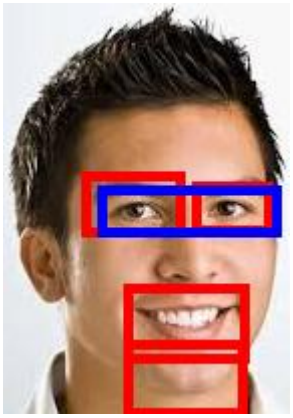
Action units refer to the muscle movement of various parts of the face. These action units are used to classify different facial expressions. This technique has been used by many researchers; however we propose the use of a modified action units classification to classify these action units and to classify different facial expressions The expressions that are classified include:

- Sadness :The outer brow and the chin are raised



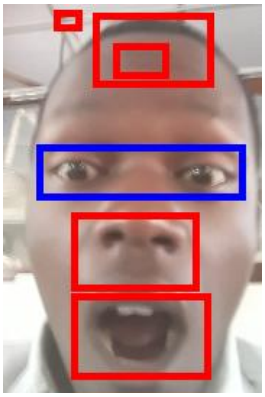
Fig(a)

- Happiness: The cheek is raised and the lip corner is pulled.



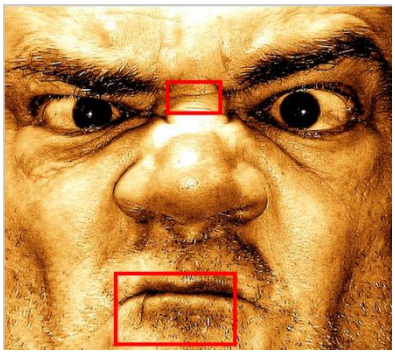
Fig(b)

- Surprise: Inner brow raised, upper lid raised and mouth is stretched



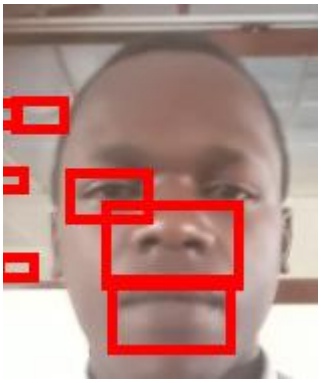
Fig(b)

- Anger: Nose wrinkle, brow lowered and lip tighten



Fig(c)

- Fear: Inner brow raised, lid tighten and brow lowered



Fig(d)

- Normal: lips part and lid tighten

Fig(e)

Step 1: Image Acquisition

The first step in facial expression recognition is image acquisition. An image is acquired using a camera and then it is stored in a database. Image acquisition in image processing can be broadly defined as the action of retrieving an image from some source, usually a hardware-based source, so it can be passed through whatever processes need to

occur afterward. Performing image acquisition in image processing is always the first step in the workflow sequence because, without an image, no processing is possible. The image that is acquired is completely unprocessed and is the result of whatever hardware was used to generate it, which can be very important in some fields to have a consistent baseline from which to work. One of the ultimate goals of this process is to have a source of input that operates within such controlled and measured guidelines that the sa me image can, if necessary, be nearly perfectly reproduced under the same conditions so anomalous factors are easier to locate and eliminate.

One of the forms of image acquisition in image processing is known as real-time image acquisition. This usually involves retrieving images from a source that is automatically capturing images. Real-time image acquisition creates a stream of files that can be automatically processed, queued for later work, or stitched into a single media format. One common technology that is used with real-time image processing is known as background image acquisition, which describes both software and hardware that can quickly preserve the images flooding into a system.

There are some advanced methods of image acquisition in image processing that actually use customized hardware. Three-dimensional (3D) image acquisition is one of these methods. This can require the use of two or more cameras that have been aligned at precisely describes points around a target, forming a sequence of images that can be aligned to create a 3D or stereoscopic scene, or to measure distances. Some satellites use 3D image acquisition techniques to build accurate models of different surfaces.

Step 2: Face Detection

Once the image has been acquired, it is subjected to a face detection process. This is to find out whether the image captured is a face or non face. Face-detection algorithms focus on the detection of frontal human faces. It is analogous to image detection in which the image of a person is matched bit by bit. Image matches with the image stores in database. Any facial feature changes in the database will invalidate the matching process.

A reliable face-detection approach based on the genetic algorithm and the eigenface technique

Firstly, the possible human eye regions are detected by testing all the valley regions in the gray-level image. Then the genetic algorithm is used to generate all the possible face regions which include the eyebrows, the iris, the nostril and the mouth corners.

Each possible face candidates is normalized to reduce lightning effect caused due to uneven illumination and the shirring effect due to head movement. The fitness value of each candidate is measured based on its projection on the Eigen-faces. After a number of iterations, all the face candidates with a high fitness value are selected for further verification. At this stage, the face symmetry is measured and the existence of the different facial features is verified for each face candidate

Step 3: Feature Extraction

The next step after detection is feature extraction. Features such as eyes, eyebrows, nose, eyelids lips and mouth are extracted. These are the features that are used in facial expression recognition.

Feature extraction a type of dimensionality reduction that efficiently represents interesting parts of an image as a compact feature vector. This approach is useful when image sizes are large and a reduced feature representation is required to quickly complete tasks such as image matching and retrieval

Step 4: Classification

Action units are used to identify different expressions. Different facial expressions are identified by analyzing the action units of different features of the face.

The table below shows different expressions and facial features used to identify them.

	Happy	Sad	Anger	Surprise	Fear	Neutral
Nose						
Eyebrows			✓	✓		
Cheek	✓					
Chin						
Lips	✓		✓	✓		✓
Lid				✓		✓
Mouth				✓		

In the table above, we can see that eight facial features are used to identify different expressions.However, the same expressions can be identified by leaving out two facial features: Cheek and chin as shown in the table below.

	Happy	Sad	Anger	Surprise	Fear	Neutral
Nose						
Eyebrows			✓	✓		
Lips	✓		✓	✓		✓
Lid				✓		✓
Mouth				✓		

From the above table, we can see that it is possible to identify six expressions using six facial features as shown below:

- Sadness: The outer brow rose.
- Happiness: The lip corner is pulled.
- Surprise: Inner brow raised, upper lid raised and mouth is stretched
- Anger: brow lowered and lip tighten
- Fear: Inner brow raised, lid tighten and brow lowered
- Normal: lips part and lid tighten

XII. CONCLUSION

This paper explores various methods of classification pertaining to facial expressions . The facial features such as eyes, lips, eyebrows,nose and mouth are used in this process. The proposed technique is efficient and reliable, hence this research work would serve as a basis for further research in this area.

REFERENCE

[1] P. S. Aleksic and A. K. Katsaggelos, Automatic Facial Expression Recognition Using Facial Animation Parameters and Multistream HMMs, IEEETransactions on Information Forensics and Security, Vol. 1, No. 1, pp.3-11, March 2006.

[2]S. M. Lajevardi and H. R. Wu, Facial Expression Recognition in Perceptual Color Space, IEEE Transactions on Image Processing, Vol. 21, No.8, pp. 3721-3732, August 2012.

[3]P. Ekman, and W. Friesen, "Facial Action Coding System: A Technique for the Measurement of Facial Movements", Consulting Psychologists Press, California, 1978

[4]Z. Hammal, L. Couvreur, A. Caplier, and M. Rombaut, "Facial expression recognition based on the belief theory: comparison with different classifiers," Image Analysis and Processing–ICIAP. Springer Berlin Heidelberg, pp. 743-752, 2005

[5] I. Kotsia and I. Pitas, Facial Expression Recognition in Image Sequences Using Geometric Deformation Features and Support Vector Machines, IEEE Transactions on Image Processing, Vol. 16, No. 1, pp.172-187, November 2007.

[6]F. Abdat, C Maaoui, and A. Pruski, "Human-computer interaction using emotion recognition from facial expression," Computer Modelling and Simulation (EMS), Fifth UK Sim European Symposium on. IEEE, pp. 196-201, 2011

[7] Y. Tian, T. Kanade and J. F. Cohn, Recognizing Action Units for Facial Expression Analysis, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, No. 2, February 2001.

[8]T. Kanade, J. F. Cohn, and Y. Tian, "Comprehensive database for facial expression analysis," Automatic Face and Gesture Recognition, 2000. Proceedings, Fourth IEEE International Conference on, pp. 46-53, IEEE, 2000.

[9]Nadir Nourain Dawoud , Brahim Belhaouari Samir , Josefina Janier ,"Fast Template Matching Method N. Sarode and S. Bhatia, *Facial Expression Recognition*, (IJCSE) International Journal on Computer Science and Engineering, Vol. 02, No. 05, 2010.Based Optimized Sum of Absolute Difference Algorithm for Face Localization" International Journal of Computer Applications (0975 – 8887) Volume 18–No.8, March 2011.

[10]. Tasnim Tarannum∗, Anwesha Pauly and Kamrul Hasan Talukder "Human Expression Recognition Based on Facial Features", 2016 5th International Conference on Informatics, Electronics and Vision (ICIEV).

[11]. J. Kalita and K. Das, *Recognition of Facial Expression Using Eigenvector Based Distributed Features and Euclidean Distance Based Decision Making Technique*, (IJACSA) International Journal of Advanced

Computer Science and Applications, Vol. 4, No. 2, 2013

[12]. N. Sarode and S. Bhatia, *Facial Expression Recognition*, (IJCSE) International Journal on Computer Science and Engineering, Vol. 02, No. 05, 2010.

[13]S. Zhou and R. Chellappa, "Probabilistic human recognition from video," in Proceedings of the European Conference on Computer Vision, pp. 681–697, Copenhagen, Denmark, 2002.

[14] V. Kr¨ueger and S. Zhou., "Exemplar-based face recognition from video.", In Proc. European Conf. on Computer Vision, volume 4, pp: 732-746.

[15] S. Zhou, V. Krueger, R. Chellappa, "Face recognition from video: A condensation approach ," in IEEE Int. Conf. on Automatic Face and Gesture Recognition, 2002, pp. 221-228.

[16] N. Vaswani and R. Chellappa, "Principal components null space analysis for image and video classification," IEEE Transactions on Image Processing, vol. 15, no. 7, pp. 1816–1830, 2006

[17] S. Soatto, G. Doretto, and Y. Wu, "Dynamic textures," in Proceedings of the International Conference on Computer Vision, vol. 2, pp. 439–446, Vancouver, Canada, 2001

[18]M. Kim, S. Kumar, V. Pavlovic, and H. Rowley, "Face tracking and recognition with visual constraints in real-world videos," in Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition (CVPR '08), June 2008.

[19] C. Shan, S. Gong, P. Mcowan, Learning gender from human gaits and faces,IEEE International Conference on Advanced Video and Signal based Surveillance, 2007, pp:505-510.

[20] Christian Micheloni, Sergio Canazza, Gian Luca Foresti; Audio-video biometric recognition for non-collaborative access granting; Visual Languages and Computing, 2009

[21] M. Balasubramanian , S. Palanivela, and V. Ramalingama; Real time face and mouth recognition using radial basis function neural networks; Expert Systems with Applications, Vol:36(3), pp: 6879-6888