IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)
e-ISSN: 2278-2834,p-ISSN: 2278-8735
PP 38-45
www.iosrjournals.org

A Neural Networks Approach for Emotion Detection in Humans

S. Kayalvizhi [1] , S. Saravana Kumar [2]
[1] *Dept. of Communication Systems , Anna University regional campus Coimbatore , India*
kayalvizhi.bava@gmail.com
[2] *Dept. Of Communication Systems, Anna university regional campus Coimbatore , India*
sskaucbe@gmail.com

***Abstract*:** *Human emotional facial expressions play an important role in interpersonal relations. This is because of humans demonstrated and conveys a lot of evident information visually rather than verbally. Although humans recognize facial expressions virtually without effort or delay, reliable expression recognition by machine remains a challenge as of today. Emotions are feeling or response to particular situation or environment. Emotions are an integral part of our existence, as one smiles to show greeting, frowns when confused, or raises one's voice when enraged. It is because we understand other emotions and react based on that expression only enriches the interactions. To automate recognition of emotional state, machines must be taught to understand facial gestures. In this paper we developed an algorithm which is used to identify the person's emotional state through facial expression such as sad, angry, happy, and normal, Existed smile, Normal smile, Confused, Fear etc. This can be done with different shape features based on face expression. According to the facial point's expression to be classified.*

Keywords: *Neutral vs emotion classification, constrained local model, key emotional points, local binary pattern, viola jones algorithm, action units.*

I. Introduction

Human beings express their emotions in everyday interactions with others. Emotions are frequently reflected on the face, in hand and body gestures, in the voice, to express our feelings or liking. Recent Psychology research has shown that the most expressive way humans display emotions is through facial expressions. Mehrabian indicated that the verbal part of a message contributes only for 7% to the effect of the message as a whole, the vocal part for 38%, while facial expressions for 55% to the effect of the speaker's message.

Emotions are feeling or response to particular situation or environment. Emotions are an integral part of our existence, as one smiles to show greeting, frowns when confused, or raises one's voice when enraged. It is because we understand other emotions and react based on that expression only enriches the interactions. Computers are "emotionally challenged". They neither recognize other emotions nor possess its own emotion.

To enrich human-computer interface from point-and-click to sense-and-feel, to develop non intrusive sensors, to develop lifelike software agents such as devices, this can express and understand emotion. Since computer systems with this capability have a wide range of applications in different research arrears, including security, law enforcement, clinic, education, psychiatry and Telecommunications.

II. Related works

2.1 Speech Recognition

Speech Recognition offers greater freedom to employ the physically handicapped in several applications like manufacturing processes, medicine and telephone network. the speech recognition system without speaker identification. how the speaker identification followed by speech recognition improves the efficiency. With this approach, the database will be divided into smaller divisions with respect to different speakers. Hence the speech recognition rate improves for the corresponding speaker.

2.2 Fingerprint Recognition

Fingerprint is the most widely used biometric system and have a very important role in forensic and civilian application, so it will continue to be used with many governments' legacy systems. The most widely and well-known used method for fingerprint is minutiae extraction method. Minutiae of fingerprint include ridge bifurcations, ridge ending, short ridge and enclosure.

Fingerprint matching techniques can be placed into two categories: minutiae-based and correlation based. Minutiae-based techniques first find minutiae points and then map their relative placement on the finger.

While correlation-based method is able to overcome some of the difficulties of the minutiae-based approach. However, it has some of its own shortcomings.

2.3 Recognition Eye Blinking

Automatic facial expression recognition has been limited to recognition of gross expression categories (e.g., joy or anger) in posed facial behaviour under well-controlled conditions (e.g., frontal pose and minimal out-of-plane head motion). We have developed a system that detects a discrete and important facial action (e.g., eye blinking) in spontaneously occurring facial behaviour that has been measured with non-frontal pose, moderate out-of-plane head motion, and occlusion.

A limitation of almost all research to date in automatic facial expression recognition is that it is limited to deliberate facial expression recorded under controlled conditions that omit significant head motion and other factors that complicate analysis. Automatic recognition of facial action units in spontaneously occurring facial behaviour presents multiple challenges.

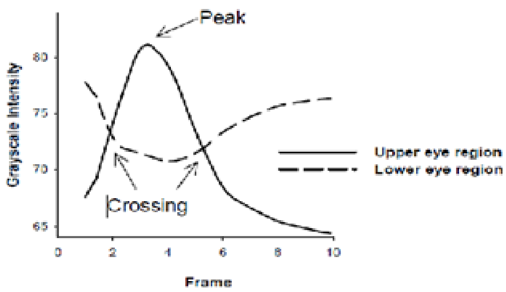

Fig 1. *Graph Based Eye Region*

2.4 Human Emotions by Recognition Gestures and Motion

The human emotion identification by gestures and movements can be useful and necessary in various areas such as communication of deaf people; education/learning; emergency services; monitoring unstable emotional state of pilots, drivers, operators of complex technical system, etc; monitoring public places to prevent illegal and extremist actions, and so on. In system works as follows, The user enters text and system displays an animated image of the gesture. A sample output of the animated gesture is shown in Fig.1.4 human gestures and motions described in a limited natural language.

III. Frame Work

In emotional classification there are two basic emotions are there, Love-fear. Based on this we classify the emotion into positive and negative emotions. The six basic emotions are angry, happy, fear, disgust, sad, surprise. One more expression is neutral. Other emotions are Embarrassments, interest, pain, shame, shy, anticipation, smile, laugh, sorrow, hunger, curiosity. In different situation of emotion, Anger may be expressed in different ways like Enraged, annoyed, anxious, irritated, resentful, miffed, upset, mad, furious, and raging. Happy may be expressed in different ways like joy, greedy, ecstatic, fulfilled, contented, glad, complete, satisfied, and pleased.

Geometry based methods track key feature points on the face, and the spatial and temporal distances between the points are used as features to classify into various facial Action Units (AUs) . Accuracy of these methods heavily relies on facial feature points fitting, which is sensitive to realistic conditions, such as facial biases, camera/head motions, lighting variations,etc. Disgust may be expressed in different ways like contempt, exhausted, peeved, upset, and bored.

In emotional expression face, If we are angry, the brows are lowered and drawn together, Vertical lines appear between the brows, lower lid tensed, eyes hard stare or bulging, lips can be pressed firmly together with corners down or square shape as if shouting, nostrils may be dilated, the lower jaw juts out. If we are happy, corners of the lips are drawn back and up, mouth may or may not be parted, teeth exposed, a wrinkle runs from outer nose to outer lip, cheeks are raised, lower lid may show wrinkles or be tense, crows feet near the outside of the eyes. Disgust has been identified as one of the basic emotions.

Its basic definition is “bad taste” secondarily to anything which causes a similar feeling, through the sense of smell, touch and even of eyesight. The well defined facial expression of disgust is characterized by furrowing of the eyebrows, closure of the eyes and pupil constriction, wrinkling of the nose, upper lip retraction and upward movement of the lower lip and chin, drawing the corners of the mouth down and back This research describes a neural network based approach for emotion classification. We learn a classifier that can recognize three basic emotions is shown in figures.

Facial Emotion Recognition has attracted significant amount of research interest. It has several applications in the areas such as humanoid robots, car industry, and various internet and mobile based services. Many supervised approaches have been investigated to detect human emotions from facial expression analysis, which are broadly classified as appearance and geometry based. Supervised methods based on support vector machines or any other similar classification strategies are trained offline using limited amount of data, which may not generalize well across people with various kinds of racial origins facial biases.

Geometry based methods track key feature points on the face, and spatial and temporal distances between the points are used as features to classify into various facial action units. Accuracy of these methods heavily relies on facial features points fitting which is sensitive to realistic conditions, such as facial biases, camera/head motions ,lighting variations. Moreover, supervised methods trained on geomentric features may suffer as the geomentric distances between the feature points in real world faces might be different to the trained ones even for neutral faces.

Appearance based methods classify textural appearance changes on the face into various facial Aus. These methods are based on supervised learning from a limited pre-defined database, and hence may not generalise well across all kinds of real world faces due to variability introduced by pose and lighting variations. As a result, it is observed that many false alarms are raised in case of unseen neural networks, thereby making neural face detection in the wild, a very challenging problem.

3.1 Constrained Local Model

The method fits constrained local model(CLM) to face, as CLMs has been proved better than the exsiting state of the art, such as active shape model, active appearance model etc., for aligning facial feature points across the frames with the help of local patch models and models and global shape constraints. As CLM is only best suited to track feature-rich points such as eyebrows, eyes, mouth etc.

Emotion textural changes happening over non-features regions such as cheek, forehead are captured with an additional set of points, named as key emotion(KE) points. The KE points by adding offsets to the CLM tracked points in the common shape for the initial frame, and later, these are tracked using temporal procrustes transformation matrix constructed between two continuous normalised CLM input shapes. The selection of KE points for change detection is based on their sensitivity to specific facial AUs.

Textural change inference at the KE points is also used to improve ER classification accuracy. Local Binary Pattern (LBP), a robust textural descriptor, is used to construct the model in our approach. LBP is proven to be invariant to monotonic illumination variations, and hence is good tracking variations in lighting conditions. Moreover, being a region based descriptor; it may also robust in case of CLM misalignments.

The contributions are summarized as follows.

1. To the best of our knowledge, it is the first time a fast online pre-processing method is proposed to detect neutral vs. any other emotion using personalised appearance models in order to take care of accuracy and speed limitations of traditional supervised methods.
2. Textural changes by CLM misalignments due to pose/illumination variations are handled by novel technical contributions such as affine noise based textural patch model generation, multi-neigbor comparison through structural similarity, and temporal transformation based KE points tracking.
3. Textural change inference at the KE points are used to improve AU classification accuracy.

3.2 KEY EMOTIONAL POINTS

3.1.1 Selection of Key Emotion Points

CLM tracked points are not affected equally by the facial expressions. For example, mouth corner CLM points are affected majorly. Lower eye brow CLM points are affected majorly. Hence consideration of all the CLM points for textural change computation would average out the changes that occur only at very few points, resulting in the missing of emotions. Moreover, there are AU informative facial regions in the face such as cheek that are not tracked by the CLM as these are not rich in texture.

Due to this, in addition to using a very few selected CLM tracked fiducial points directly for change detection, a few KE points are generated over eyebrow and cheek with respect to a set of closest stable CLM tracked points in the common space by adding fixed offsets at $t=0$. Eye corner CLM points are used to generate eyebrow KE points. Cheek KE points are generated using vertical co-ordinate of CLM nose and horizontal coordinate of CLM eye points.

The KE points chosen at the mouth corners are sensitive to AUs at the cheek are sensitive to AU, and at the eyebrow are sensitive to Aus. An example for showcasing the sensitivity of KE points for various AUs. One may observe that a huge change in Mahalanobis distance values at the KE points for emotion frames than the neutral frame on comparing with the corresponding models.

3.1.2. Tracking of KE points

As stated KE points are generated in the reference shape by adding fixed offsets to a set of stable CLM points. However, these fixed offsets may not be consistent with time due to CLM alignment errors caused by pose/expression variations, resulting inaccurate KE points.

To alleviate this, KE points generated at *t=0*, where the chances of CLM fitting accuracy are high, is tracked in the later frame using a sequence of transformations, current normalized shape is used for accurate tracking of KE points, Transform parameters are calculated using non-emotive CLM points.

For t=0 *Step.1*. KE points are derived using offsets over stable CLM points in the reference shape. 1 *tm* is computed between reference and current shapes. KE points in the reference shape are mapped into current shape using 1 *tm* .

For (t>0) *Step.2*. By aligning current shape to the previous normalized shape using 2 *tm* , current normalized shape, devoid of affine variations between two shapes is obtained. “Previous shape” is considered for previous normalized shap

For t=1. *Step.3*. As the variations between two continuous normalized input shapes are gradual in general, fiducial point tracking is accurate. Hence the transformation matrix between two continuous normalized input shapes is also accurate. 3 *tm* is computed between previous and current normalized shapes.

Step4. KE points in the previous normalized shape are mapped to current normalized shape by 3 *tm* .

Step.5. 4 *tm* is computed between current normalized shape and aligned/reference shape.

Step.6. As patch extraction is done in the common shape, KE points in the current normalized shape are mapped to the common shape (reference/aligned shape) based on 4 *tm* .These four transformations are required for accurate mapping of KE points from the input space to the common space.

In fusion distance KE points are fused separately for each region-mouth ,cheek, eyebrow. The method of fusing the distances in each symmetrical half of the face is critical in deciding the overall accuracy. The blocks are described in proposed method in fig 2. If the KE points are asymmetrical across both the halves of the face due to CLM misalignments. The maximum of points distances is chosen and threshold to identify the changes in the locations of the points in the pair.

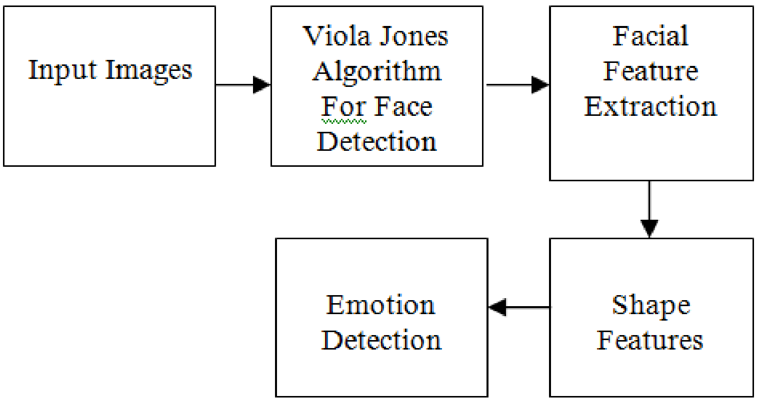

Fig 2. *Block Diagram*

3.3. Viola Jones Algorithm

The basic principle of the Viola-Jones algorithm is to scan a sub-window capable of detecting faces across a given input image. The standard image processing approach would be to rescale the input image to different sizes and then run the fixed size detector through these images. This approach turns out to be rather time consuming due to the calculation of the different size images.

Contrary to the standard approach Viola-Jones rescale the detector instead of the input image and run the detector many times through the image – each time with a different size. At first one might suspect both approaches to be equally time consuming, but Viola-Jones have devised a scale invariant detector that requires the same number of calculations whatever the size. This detector is constructed using a so-called integral image and some simple rectangular features reminiscent of Haar wavelets. The next section elaborates on this detector.

3.4 Facial Feature Extraction

Applying human visual property in the recognition of faces, people can identify face from very far distance, even the details are vague. That means the symmetry characteristic is enough to be recognized. Human face is made up of eyes, nose, mouth and chin etc. There are differences in shape, size and structure of those organs, so the faces are differ in thousands ways, and we can describe them with the shape and structure of the organs so as to recognize them.

One common method is to extract the shape of the eyes, nose, mouth and chin, and then distinguish the faces by distance and scale of those organs in fig 3.6.The other method is to use deformable model to describe the shape of the organs on face subtly.

We can tell the characteristics of the organs easily by locating the feature points from a face image. If we normalized the characteristics which have the properties of scale, translation and rotation invariance, we can normalize the faces in the database through pre-treatment, so as to extend the range of database, reduce the storage and recognize the faces more effectively. The input image can be taken in fig 3

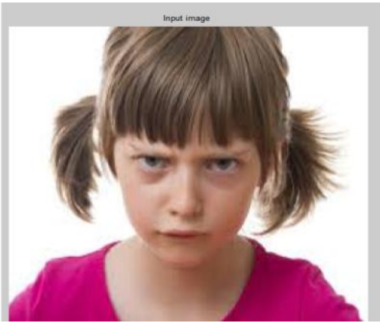

Fig 3. *Facial Action Units (Aus)*

Additionally, the selection of face feature points is crucial to the face recognition. The number of the feature points should take enough information and not be too many. If the database has different postures of each people to be recognized, the property of angle invariance of the geometry characteristic is very important.

A method to locate the vital feature points of face, which select 4 feature points that have the property of angle invariance, including 2 eyeballs, the midpoint of nostrils and mouth. According to these, we can get other feature points extended by them and the characteristics of face organs which are related and useful to face recognition.

3.4.1 Geometric Features

Measure similarity bet, Shapes by measuring similar bet. Their features are Generally, simple geom. features cannot discriminate shapes with large distances e.g. rectangle vs. Ellipse .

• Usual combine with other complimentary shape descriptors and also used to avoid false hits in image retrieval

• Shapes can be described in fig 4 many aspects we call shape parameters: center of gravity/centroid, axis of least inertia, digital bending energy, eccentricity, circularity ratios, elliptic variance, rectangularity, convexity, solidity, Euler number, profiles, and hole area ratio.

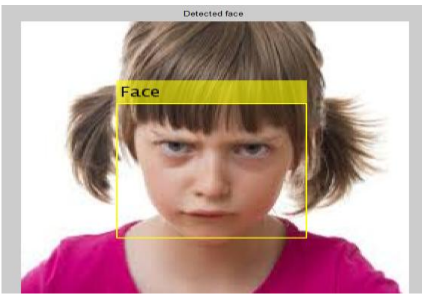

Fig 4. *Rectangular GeomenticFeatures*

3.4.2 Shape Features:

Identability: shapes which are found perceptually similar by human have the same features that are different from the others. Translation, rotation and scale invariance: the location, the rotation and the scaling changing of the shape must not affect the extracted features. Affine invariance: the affine transform performs a linear mapping from coordinates system to other coordinates system that preserves the "straightness" and "parallelism" of lines. Affine transform can be constructed using sequences of translations, scales, flips, rotations and shears.

The extracted features must be as invariant as possible with affine transforms. Noise resistance: features must be as robust as possible against noise, i.e., they must be the same whichever be the strength of the noise in a give range that affects the pattern.Occultation invariance: when some parts of a shape are occulted by other objects, the feature of the remaining part must not change compared to the original shape.statistically independent: two features must be statistically independent. This represents compactness of the representation. Reliability: as long as one deals with the same pattern, the extracted features must remain the same .

3.5 Analysis Of Emotion Detection Using Artificial Neural Network

The image processing section was built using MATLAB functions and it comprises techniques such as feature extraction, where facial features and geometric features are extracted to analyze the emotions angry, sad, fear, neural, happy. The Back propagation Neural Network was used in this project to enhance the accuracy and performance of the image processing

3.5.1 Software Simulator

MATLAB is a high-performance language for technical computing. It integrates computation, visualization, and programming in an easy-to-use environment where problems and solutions are expressed in familiar mathematical notation.

Typical uses include the graphical output is optimized for interaction. You can plot your data very easily, and then change colors, sizes, scales, etc, by using the graphical interactive tools. MATLAB's functionality can be greatly expanded by the addition of toolboxes. These are sets of specific functions that provided more specialized functionality .Example: Excel link allows data to be written in a format recognized by Excel, Statistics Toolbox allows more specialized statistical manipulation of data (ANOVAs, Basic Fits, etc).

IV. Results And Discussion

A method to locate the vital feature points of face, which select 4 feature points that have the property of angle invariance, including 2 eyeballs, the midpoint of nostrils and mouth.
Extract key emotional points as follows : A Constrained Local Model (CLM) is class of methods of locating sets of points (constrained by a statistical shape model) on a target image.
1. Align current shape to previous normalized shape.
2.Compute difference between two continuous normalized input shapes.
3. Map KE points.
4. Computed tm4. Tm4 is computed between normalized shape and aligned/reference shape.
5. Extract patch is shown in fig 5.

1. Lip Feature: Lip is a sensory organ existing in visible portion of human face and considered to be different for each and every individual. There are researches carried out for face recognition and classification of gender using lip shape and color analysis.

2. Eye Feature: Eye has randomness due to its small tissues which provides differentiation to the pattern of eye for each and every individual human being. The stableness, uniqueness and non-invasion, these qualities make the iris outstanding among several biometric feature.

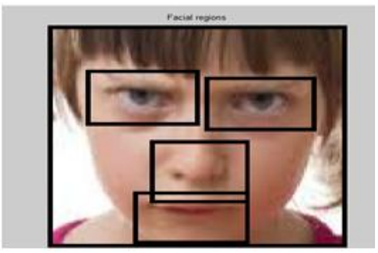

Fig 5. *Detection Of Facial Features*

4 **Nose Feature**: The nose tip is a distinctive point of human face. It also remains unaffected even due to changes in facial expressions. Thus, it proves to be efficient for face recognition.

Fig 5. *Eyes.Nose And Mouth Feature Extraction*

5 **Image Recognition** :Recognition is the process that assigns a label to an object based on the information provided by its descriptors. Interpretation involves assigning meaning to an ensemble of recognized objects. Feature extraction and Emotion detection as shown in figure.

Fig 6. *Emotion Detection*

For example, identifying a character as, say, c requires associating the descriptors for that character with the label c. Interpretation attempts to assign meaning to a set of labeled entities. For example, a string of five numbers are followed by a hyphen and four more numbers can be interpreted to be a ZIP code.

V. Conclusion

The working of various blocks in this project has been explained in detail in the above chapters. The experimental results and comparison is also done in this chapter. From the results it is proved that the performance of the back propagation of artificial neural network is better compared to other existing methods. The development of a program which determines the emotions present in face through image processing coupled with artificial neural network was successfully implemented using MATLAB. Also, typical lesions of the emotions were successfully identified by having an accuracy of 100% for the emotion detection. In future work of this project is to add more type of emotions that can be detect using the program and To collect more images for the dataset of neural network which could yield more accurate results.

References

[1]. Bartlett M.S, Littlewort G and Frank M.G Etl (2005) “Recognizing facial expression: Machine learning and spontaneous behaviour” in proc.IEEE CVPR, pp.568-573.

[2]. Cristinacce D and Cootes T.F (2006) “Feature detection and tracking with constrained local models” In BMNC, pp.929-938.

[3]. Cohn J.F, Xiao J and Moriyama T Etl (2003) “Automatic recognition of eye blinking in spontaneous facial behaviour” in Behaviour Research Mds, Inst.and Compu., Vol35,pp.420-428.

[4]. Kotsia I and Pitas I (2007) “Facial expression recognition in image sequences using geometric deformation features and support vector machines,”in proc.of IEEE Trans.6, Image Process.,Vol. no.1,pp.172-187.

[5]. Pantic M and Rothkrantz L (2004) “Facial action recognition for facial expression analysis from static face images,” In IEEE Trans.Syst., Man,Cybern.B.vol.34, no.3,pp.1449-1461.

[6]. Shan C, Gong S and McOwan PW (2009) “Facial expression recognition based on local binary patterns: A comprehensive study” Image and Vision Comput.,vol.27,no.6,pp.803-816.

[7]. Tian Y, Kanade T and Cohn J (2001) "Recognizing action units for facial expression analysis" in proc.of IEEE Trans.PAMI., vol.23,no.2.pp.97-115.
[8]. Tong Y, Liao W and Ji Q (2007) "Facial action unit recognition by explotting their dynamic and semantic relationships"in Proc. .29,no.10,pp.1683-1699.
[9]. Wang Y, Lucey S and Cohn J.F,(2008) "Enforcing convexity for improved alignment with constrained local models," in IEEE Conf.CVPR,pp1-8.
[10]. Zeng Z, Pantic M, and Huang T (2009) "A survey of affect recognition methods: Audio, visual and spontaneous expressions,"in IEEE Trans.PAMI.,vol.31,no.1,pp.39-58.