Facial Expression Recognition using Fusion of LBP and HoG Features

Mamta Santosh, Avinash Sharma

Abstract**: *Human Machine Interaction has gained significant attention in the previous years. The communication through facial expression plays a vital role in the interaction since the expressions convey more information than spoken words and it allows to recognize the emotions without wearing sensors and monitoring devices. So it is important to classify the emotions correctly. In this paper, we have proposed an effective framework for Emotion Recognition using Facial Expressions which recognizes 96.2% of emotions accurately. Voila Jones Algorithm is used for face detection followed by splitting the face in parts which are dominant for expressions and applying Gauss Filtering to these parts. Combination of HOG and LBP features is used to extract the features. PCA is used for dimensionality reduction and these features are classified using Canberra Distance Classifier. The framework performs well on JAFFE Dataset and gives promising results.***

Index Terms**: *Emotion Recognition, Facial Expression Recognition, LBP, HOG, Canberra Distance Classifier***

I. INTRODUCTION

Emotional aspect broadly impacts on social intelligence like communication; decision making and even helps in understand the behavior of an individual. It plays essential role throughout communication. Emotion recognition is carried out in various ways, which can be verbal or non-verbal.Voice is verbal sort of communication and facial features, action, gestures is non-verbal form of communication. As per Mehrabian, in human communication only 7% communication is accounted verbally, 38% by voice intonations whereas, 55% of message is affected by facial expressions. Therefore robotic and real time facial features would play necessary role in human and machine interface.For humans it is very significant to do visual communication. The human face is involved in a large variety of different activities. Six basic emotions presented by Ekman which are claiming to be linked with different facial features. Those 6 essential emotions are: Sadness, fear, surprise, happiness, anger and disgust.

Moreover, emotions also exceed cultural diversities and quality. Face expressions helps deliver the message that is tough for others to understand. It shows the physical efforts or the intentions of an individual that he should be applying for practicing any specific task. So, it is important to recognize the emotions in type of areas like image process, cyber security, psychological studies, robotics, and video game applications to name some.

Revised Manuscript Received on June 7, 2019

Mamta Santosh, Computer Science &Engg, MMDU/MMCE, Ambala, India.

Dr. Avinash Sharma, Computer Science &Engg, MMDU/MMCE, Ambala, India.

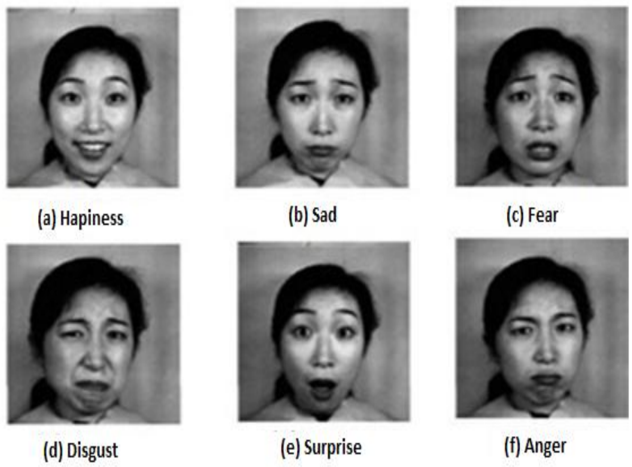


Figure 1: Basic Six Facial Expressions (JAFFE Dataset)

Efforts in the interaction domain are being created to assemble high-quality info that helps meeting the demands of the system in order that it will stimulate human emotions accurately. Hence, building a system that's capable of emotion recognition based on expressions has been an important area of analysis in present.

Various Applications of Facial Expression Recognition are:

- Entertainment
- Effective Computer Tutor
- Intelligent Customer Services
- Intelligent Home Robotics
- Gaming
- Remote Patient Monitoring
- Marketing
- Automated Detection of Driver Fatigue
- Smart Homes

Detection of face determines the sizes and locations of faces in an image. They're simply placed in untidy scenes by adults and young kids. Automatic detection of faces by computers could be terribly difficult task because face patterns will have perceptibly changeable image appearances. For instance, faces of humans vary from ages, genders, races and hairstyles etc. Additionally, the distinction of poses of faces and scales shapes in pictures also delay the success of automatic face detection systems. Some completely different approaches are practiced to unravel the matter of face detection. Every approach that is used has its own advantages and drawbacks.

The facial expressions are result of facial muscle

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

movements. To measure these facial muscle movements FACS (Facial Action Coding System) is defined. These movements are called Action units. The combination of these Action Units forms different expressions.

The proposed framework is applied on JAFFE (The Japanese Female Facial Expression) Dataset which consists of 213 images of seven emotional states. Each emotional state contains more than thirty 256×256 images of different Japanese people. It has six basic expression and neutral faces. The dataset has posed expressions [2].

II. LITERATURE REVIEW

Several techniques have been proposed in the literature for recognition of facial expression of JAFFE dataset using different frameworks.

The general approach consists of Emotion Recognition consists of these three steps:

1) Face Acquisition
2) Facial Data Extraction and
3) Facial Expression Recognition

Modified LBP (Local Binary Pattern) technique has been proposed by Yeshudas et al. [8]. They used LBP for extraction of features and for classification they used neural networks. They used two datasets for their methods. They are Japanese and Taiwanian database. Their algorithm performed better than many approaches.

Happy et al [6] presented a hybrid approach in which they combined shape and appearance features. They combined these features to form a hybrid feature vector. They extracted Pyramid of Histogram of Gradients (PHOG) as shape descriptors and Local Binary Patterns (LBP) as appearance features. They used linear discriminant analysis to reduce the dimensionality of the feature. Then Support Vector Machine classifier was used.

Local Binary Pattern (LBP) which are texture features. Different variants of LBP are used for Emotion Recognition using Facial Expression Recognition. Zhang et al. [4] proposed Neural Network-based and multi-class SVM based classifiers. The Multiclass SVM which uses radial basis function kernel enabled the robot to perform better than the Neural Network-based emotion recognizer. They used 380 training samples to train NN based classifier and used 477 test examples to test the SVM classifiersand performed with 76% accuracy.Caifeng et al. [3] used different variants of LBP used in FER systems such as s Support Vector machine, Template Matching, Linear Discriminant Analysis, LBP extracted features using different databases.

P. P. Ghadekar et al [7] proposed Local Directional Number Pattern (LDNP) technique for feature extraction. They used Principal Component Analysis for dimensionality reduction and achieved accuracy of 82.2% on CK+ Dataset.

In literature one would often declare that the motion in the form of optical flow is one of the leading cues in the human visual system. In this estimation author uses descriptors in the form of jets in Lucas Kanade setting. Large displacements are difficult to estimate because of the linearization involved in this method.Motion estimation using correlation transform is based on the assumption that zero crossed normal cross correlation improves the accuracy of optical flow. They define very unfair descriptors based on correlation transform and use robust penalty to control motion discontinuities.

Sparse flow establishes sparse pixel resembles which are then applied to a variation approach to obtain a refined optical flow

Happy et al. [6] proposed approach that combined shape and appearance features. They used Pyramid of Histogram of Gradients as shape descriptor and used Local Binary Patters as appearance feature. They processed the active patches rather than processing the whole face. It reduced computational cost. They used Linear Discriminant Analysis for dimensionality reduction and SVM for classification. Their method achieved 94.63% recognition rate on CK+ and 83.86% on JAFFE database.

Drira et al., [1] proposed framework for analyzing, comparing, matching 3D faces and averaging their shapes. For analyzing the shapes of full facial surfaces, they used Riemannian framework to solve the challenges such as open mouths, missing parts, and partial occlusions due to hair, glasses, and large facial expressions with large pose variations. The framework used PCA for formal statistical implications which includes assessment of missing facial parts on tangent spaces and computing average shapes.

In geometric feature, firstly key facial points are detected and tracked to recognize the expression. Ghimire et al. [7] detected 52 facial key points and then they modeled these 52 facial key points in the form of points and lines feature. They used multicastAdaboost and SVM classifier for recognition of facial expressions.

III. PROPOSED FRAMEWORK

In proposed framework, Viola-Jones method and Face Landmark Detection method are used for face detection. The detected face is splitted into three parts and Gaussian filter is applied to the parts. Histogram of oriented gradients (HOG) and LBP (Local Binary Pattern)are used for feature extraction due to its superiority over other methods. To reduce the dimensionality of features Principal Component Analysis is used so that the maximum variation is preserved. Canberra distance is used for classification into different expressions.

The proposed framework is applied on most popular database i.e. JAFFE for facial expressions

A. Face detection and centering

For face detection and centering, twomethods are used: Viola-Jones method and Face Landmark Detection method. Face in image is detected by Viola-Jones [5]. This method performs well to detect frontal face or tilted face with angle maximum 10 degree. But the disadvantage with Voila Jones method is thatthe method also includes hair and some part from image that we don’t need for facial recognition. To refine the detection result, face landmark detection is applied and finally use the landmark coordinates to center face after cropping. Facial landmarks are used to localize and represent regions of the face, such as:Eyes,,Eyebrows, Nose, Mouth and Jawline. This method is as follows:

- A training set of labeled facial landmarks on an image. These images are *manually labeled,*

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

specifying **specific** *(x, y)*-coordinates of regions surrounding each facial structure.

- Given this training data, an ensemble of regression trees are trained to estimate the facial landmark positions directly from the pixel intensities.
- The pre-trained facial landmark detector inside Matlab shape_predictor_68_face_landmarks.dat is used to estimate the location of *68 (x, y)-coordinates* that map to facial structures on the face.
- The face is centered based on these points taking nose as center point.

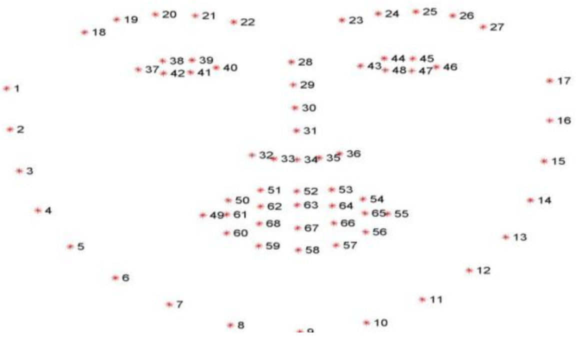


Figure 2: Face Landmark Detection

- Centered face is fed up to HOG feature extraction method.
- HOG is an edge orientation histograms based on the orientation of the gradient in localized region which are called cells.

Therefore, it is easy to express the rough shape of the object and is robust to variations in geometry and illumination changes.

Result of both process are shown below:

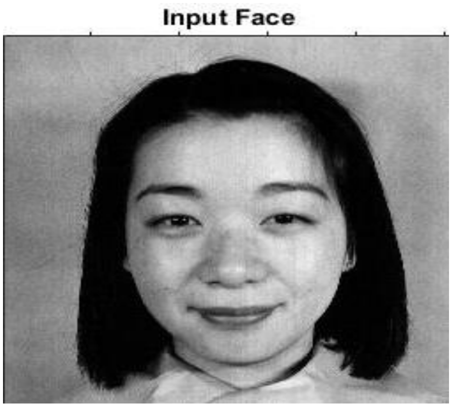


Figure 3(a): Original Face Image

Figure 4: Cropped Face with Voila Jones

Figure 5: Refinement and centering

As we can see in from figure 3-5, the detected face is refined and centered very well. Landmark face detection method operates by matching face with bunches of face model (points that surround face shape). It means that the centering will be accurate because center point is regarded to these points instead of rectangular coordinated (points) that we get from Viola-Jones. We keep using Viola Jones at the first step to speed up the landmark face detection becausedetection will be conducted only on detected face. Also, this will avoid the false detection of landmark face detection method. After cropping, face is resized to 101 x 81 pixels.

.

B. Splitting face into three parts

Face is split into 3 parts which are Eyes, Left and Right Cheek that includes Nose, Mouth. The aim of this action is to get feature of face locally. In expression recognition problem, part of faces like eyes, mouth, nose and cheek change in size, form. This fact can be used to form a hypothesis that locally extracted feature might be able to distinguish emotion better if we compare to the whole face. This also guarantees that we can extract more features from each area that can be used as one on one comparison. Below images are the samples of split area of face into 3 parts:

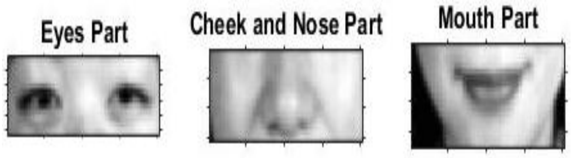


Figure 6: Split areas of face

C. Gauss Filtering

After split face extracted from a face, a Gaussian filter is applied to each of the split area. The aim is to smooth face a little bit so only strongest texture that we can extract. Some small texture like wrinkles can be suppressed to reduce false negative. Parameters that we set for Gaussian filter is sigma = 0.95 and hsize = [5 5]. In our method, changing the value of sigma and hsize bigger or smaller impacted the final result. Accuracy went lower compare to our current observation parameters value.

D. Facial Feature Extraction

In this research, we apply two feature extraction methods to get two different types of features. They are HOG and LBP. Both methods are good at detection of texture feature on face.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

1) HOG Features

HOG is purely dependent on the localized region, that is called cells, and as that of the gradient it is dependent on its orientation. Moreover, it Histogram of oriented gradients (HOG) is based on the orientation of the gradient in localized region known as cells. Therefore, it’s very simple process to convey the rough form of the object and is healthy to a variation that takes place in illumination changes and geometry. On the other hand, rotation and change of scales are not supported. Generally, HOG method is explain by these steps:

- The main pictures are converted into grayscale.
- At each pixel, gradient is calculated.
- To create a bar graph (Histogram) of gradient orientation for every cell using gradient magnitude and orientation, size of cell that is used is 9 x 9 with number of 9 bins.
- Normalization and Descriptor blocks

HOG features of these areas are visualized by figure below

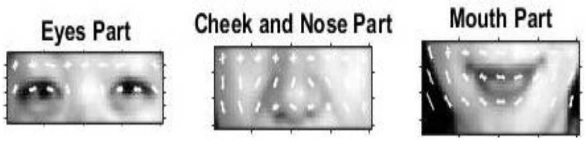


Figure 7: HoG Features

Solid white lines over the face describe the HOG features for current image. Extracted feature has size 1 x 250 for each of area that makes the total become 750 for three areas. This feature will be concatenated with LBP features.

2) LBP Features

There are many ways of extracting the most helpful features to perform facial recognition. LBP (Local Binary Pattern) labels the pixels of a picture by a very easy process and is also an efficient texture operator. There are a number of features so far and among them one is called the LBP technique. Using Local Binary Pattern, it’s attainable to explain the form of a digital image and its texture. The features can be extracted by dividing the pictures into many little regions. Because of its discriminative power and ease in computations, Local Binary Pattern the operator of texture has become a well-liked approach in numerous applications. Among all these, one of the mostly known and popular feature of LBP has been the strength it has for the monotonic changes it holds for the gray-scale.

Perhaps the most significant feature of the Local Binary Pattern operator in real-world is its strength to monotonic changes in gray-scale, as computational simplicity is one of its significant properties which in challenging real-times make it possible to analyze pictures.

Local Binary Pattern is practiced to each split area of face. Value of some parameters has been decided based on experiment. They are radius = 3 and cell size is [9 9]. Each of area gives 240 features that make them 720 at total.

.

E. Dimensionality Reduction using PCA

In the proposed method, the number of images for training is 70 images that come from 10 different persons. Each person contributes 1 image for each of expression. This makes the total features is 1470 x 70. By using PCA, this dimension was reduced into 70 x 70 and used as template.

F. Classification

This part also use face detection and centering and feature extraction step as explained in the Training Part. The final step of this part is operating PCA parameters that we get in training part. Parameters from PCA that we use are image mean and its Eigen vector. Image mean is used to subtract feature of input image and Eigen vector is multiplied to the result to get its weight. This weight will be used as input of distance calculation with the weight from training datasets.

The distance calculation is based on Canberra Distance as shown below

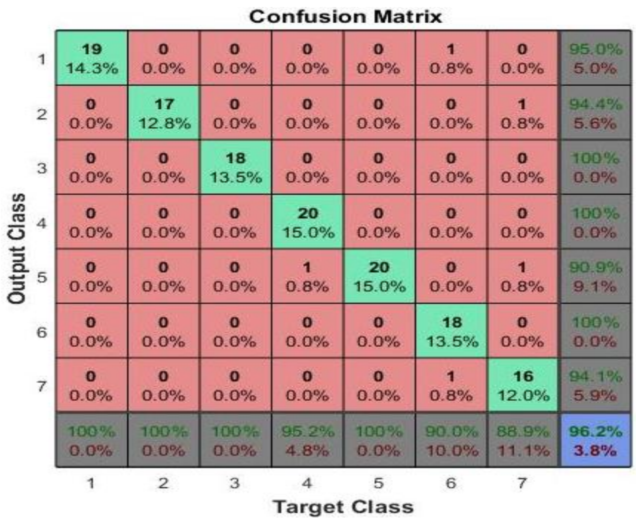
$$d^{CAD}(i,j) = \sum_{k=0}^{n-1} \frac{|y_{i,k} - y_{j,k}|}{|y_{i,k}| + |y_{j,k}|}$$

.

The minimum distance is the winner or recognized emotion. The accuracy of recognition is shown with confusion matrix. In this testing, I use datasets from JAFFE database. The emotion to be recognized is 7 emotions. I use 70 images as training and use 133 images for testing. The training and testing images are different. Beside the accuracy calculation, I also plot the confusion matrix.

An accuracy of 96.2 % is achieved for six basic expressions and neutral face. Confusion matrix is as shown below:

1. Anger 2. Disgust 3. Fear 4. Happy 5. Neutral 6.Sad 7.Surprise The confusion matrix for 7 emotions is



Output \ Target	1	2	3	4	5	6	7	
1	19 14.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 0.8%	0 0.0%	95.0% 5.0%
2	0 0.0%	17 12.8%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 0.8%	94.4% 5.6%
3	0 0.0%	0 0.0%	18 13.5%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
4	0 0.0%	0 0.0%	0 0.0%	20 15.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
5	0 0.0%	0 0.0%	0 0.0%	1 0.8%	20 15.0%	0 0.0%	1 0.8%	90.9% 9.1%
6	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	18 13.5%	0 0.0%	100% 0.0%
7	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 0.8%	16 12.0%	94.1% 5.9%
	100% 0.0%	100% 0.0%	100% 0.0%	95.2% 4.8%	100% 0.0%	90.0% 10.0%	88.9% 11.1%	96.2% 3.8%

Figure 8: Confusion Matrix

IV. CONCLUSION

In this paper, an effective emotion recognition framework using facial expression is proposed in which Viola Jones and Landmarkdetection methods are used for face detection. The detected face is splitted into three parts which are dominant for the movement of face and cause the expression. Gaussian Filters are used to smooth the faces and remove false factors. HOG(Histogram of Gradients) and LBP (Local Binary Pattern) features are used to extract the features and PCA is applied fordimensionalityreduction of the features. Emotions are classified usingCanberra

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Distance Classifier. An accuracy of 96.2 % isachieved which is higher than many state of the arttechniques.JAFFE dataset is used for the training and testing of images.

ACKNOWLEDGMENT

We are thankful to Michael J. Lyons, Shigeru Akamatsu, and MiyukiKamachi&JiroGyoba for providing us the JAFFE Database.

We are thankful to University Grand Commission for providing Junior Research Fellowship for carrying out research work.

REFERENCES

1. Drira, Hassen, Boulbaba Ben Amor, Anuj Srivastava, Mohamed Daoudi, and Rim Slama. "3D face recognition under expressions, occlusions, and pose variations." IEEE Transactions on Pattern Analysis and Machine Intelligence35, no. 9 (2013): 2270-2283.
2. *Coding Facial Expressions with Gabor Wavelets*Michael J. Lyons, Shigeru Akamatsu, Miyuki Kamachi&JiroGyoba Proceedings, Third IEEE International Conference on Automatic Face and Gesture Recognition, April 14-16 1998, Nara Japan, IEEE Computer Society, pp. 200-205.
3. 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.
4. Li Zhang; Hossain, A.; Ming Jiang, "Intelligent Facial Action and emotion recognition for humanoid robots," in Neural Networks (IJCNN), 2014 International Joint Conference on , vol., no., pp.739-746, 6-11 July 2014.
5. P. Viola and M. J. Jones, “Rapid Object Detection using a Boosted Cascade of Simple Features,” in Proceedings of the IEEE Computer Society International Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 511- 518, 2001.
6. Happy, S.L.; Routray, A., "Robust facial expression classification using shape and appearance features," in Advances in Pattern Recognition (ICAPR), 2015 Eighth International Conference on , vol., no., pp.1- 5, 4-7 Jan. 2015
7. P. P. Ghadekar, H. A. Alrikabi and N. B. Chopade, "Efficient face and facial expression recognition model," 2016 International Conference on Computing Communication Control and automation (ICCUBEA), Pune, 2016, pp. 1-8.
8. Y. Muttu and H. G. Virani, "Effective face detection, feature extraction & neural network based approaches for facial expression recognition," 2015 International Conference on Information Processing (ICIP), Pune, 2015, pp. 102-107.

AUTHORS PROFILE

to

Mamta Santosh has completed her M. Tech. in Computer Science & Engineering from DCRUSTM Murthal. She has been pursuing PhD as Senior Research Fellow from Maharishi Markandeshwar ngineeringcollege,Mullana,Ambala (Haryana) since 2014. She has qualified various national level exams in the field of Computer Science such as UGC-NET-JRF 2014, GATE 2015, HTET 2014 & 2019. She has published four papers on the topic.

Presently Principal & Professor, *Maharishi Markandeshwar* Engineering college,Mullana,Ambala (Haryana) Constituent institution of *Maharishi Markandeshwar University*, Mullana is **NAAC accredited 'A' grade deemed university**& the first private engineering college of Haryana established in 1995.Also Dean Faculty of Engineering and Technology & Member of Board of Studies & DRC Committee for Research.**Ex-Principal & Professor, Rajasthan College of Engineering for Women (Leading Women’s Engineering College in the state of Rajasthan).** Publications: International Journals: Published: 45 (Accepted: 61); International Conferences: Published: 90 (Accepted: 102) National Conferences & Workshops: 75 (Accepted: 110 & more) Text Books/**EDITED**: 06.
Total Experience: 20 years (10 years PG) + 03 years Research
International Conferences Organized: 10. He is an active member of advisory/technical program committee of reputed International/National conferences & reviewer of number of reputed Journals e.g. Springer, Elsevier Journal Computers & Electrical Engineering. Approximate 20 years of rich experience in Teaching, research and industry managing technical institution, serving in all capacity including Head of Department, Professor, Controller of Examination, Dean Academics Affairs, Principal etc. Played leading role in accreditation of the institution and ISO 9001:2000 certification. (Including05 years of industrial/research experience)