Facial Expression Recognition using HOG, DWT and HC method, SVM Classifier

Ankita Chaurasiya[1], Vinod Todwal[2],Subhash Chandra Jat[3]
Department of Information Technology, Rajasthan College of Engineering for Women, Jaipur, India

Abstract : Facial expressions are the fastest means of communication to convey any information. They not only show the feelings of particulars but can also be utilized to determine their mental views. Facial expression recognition technique detects the emotion of particular by his/her facial expression. FER plays a very important role in Human Computer Interaction and Security monitoring. This requires very fast processing and decision making so we have to develop innovative recognition system that can find facial expression in an effective manner. It is very easy for a human to detect facial expression but for a machine it is a very tough task. Facial expression is most frequently used nonverbal mechanism by human's beings to dispatch internal emotional states and, therefore, plays a fundamental role in interpersonal interactions. Although there is a wide range of possible facial expressions, psychologists have identified seven universally recognized fundamental expressions (angry, disgust, fear, happy, neutral, sad and surprise). It is obvious that a system capable of performing automatic recognition of human emotions is a desirable task for a set of applications such as human-machine interaction, security, affective computing, and so on.

The purpose of this research work is to examine the fusion technique of Harris Corner, HOG and DWT features for facial expression recognition. SVM Classifier computes the similarity score by utilizing several available public databases of facial expressions JAFFE, CK and KDEF. Confusion matrix plot is used to measure the productivity of system by means of accuracy, precision and sensitivity graphs**.**

Index Terms **- Confusion Matrix Plot, Discrete Wavelet Transform, Harris Corner, Histogram of Oriented Gradients, Support Vector Machine, Viola Jones Method.**

I. INTRODUCTION

The technique of spotting individual by using particular's physiological and behavior features is known as biometric recognition some patterns which are used by biometric system to identify the individual are face of a person, fingerprint of any person, hand geometry, iris, DNA and signature. Biometric has attained so much curiosity among the researcher to identify or classify any human because of the high miscellany in features used in it.

There are numerous distinct points available on face which can be used in an algorithm or combination of algorithms for classifying the features. It is not necessary for an individual to be near the face biometrics because it is able to execute the algorithm by taking the input image of face.

In various security application this system is considerable like forensic department, intelligence department and government crucial security department. Behavioral or physiological traits which having capability of identifying the individual makes biometric self-sufficient system. The system can use combination or mixture of pattern to provide best results pattern like face, iris, retina, fingerprint, hand geometry etc.

The biometric come up with the following characteristics which are:

- Universal
- Unique
- Permanent
- Collective
- Performer
- Acceptable
- Genuine [1]

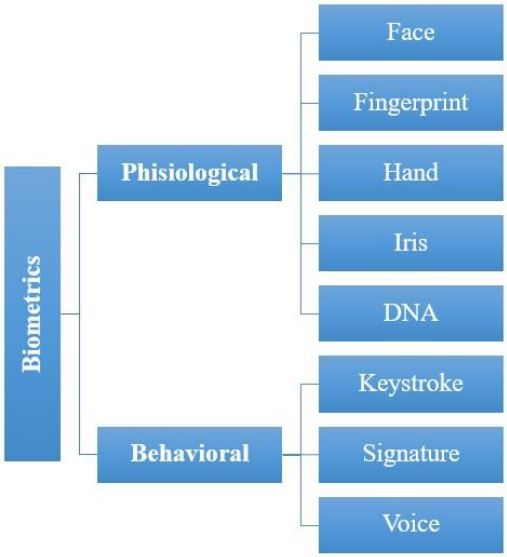

Fig. 1: Different modalities [5]

II. Facial Expression

Facial expressions are a form of nonverbal communication that allows an individual to quickly pass the information to those around him about his or her own emotional state. It turns out that the human being is particularly adequate to recognize the emotions expressed by others on his face. Thus, the paradigm of recognition of emotional facial expressions has been the subject of many researches because it is a preferred way to study emotions. Researchers in neuroscience and psychology, however, consider emotion as a complex phenomenon and the study of the recognition of facial expressions cannot be conceived without exposing, beforehand, basic notions about emotions.

.

III. Facial Expression Databases

One of the most important aspects of the development of any new facial expression recognition or detection system is the choice of the database that will be used to test the new system or method because the quality cannot be assured for real time input image, the accuracy provided by the algorithm created is expected to be high [2]. That's why common databases are needed to evaluate the algorithms in a comparative way. In this section, we are going to present some of the databases of popular facial expressions that are publicly and freely available. Other databases are available and their coverage is not done here.

3.1 Cohn-Kanade (CK) Database:

The CK database [6] is very popular and has been widely used by the facial expression recognition community for testing the system. The CK database contains 97 university students aged between 18 and 30 years old.

Fig 2: Examples of images extracted from the CK database

3.2 Karolinska Directed Emotional Faces (KDEF) Database

The KDEF database [7] was originally developed for neuroscience research. However, now it has since been widely used in the field of computer vision because of its applicability.

Fig. 3: Examples of images extracted from the KDEF database

3.3 Japanese Female Facial Expression (JAFFE) Database

The JAFFE database [7] contains 213 images of facial expression of ten Japanese women. The latter simulated 3 to 4 examples for each of the six basic emotions, as well as neutral emotion. The resolution of the images is 256×256 pixels.

Fig. 4: Examples of images extracted from the JAFFE database

Expressions of database from left to right: Neutral, Joy, Sadness, Surprise, Anger, Fear, Disgust.

IV. PROPOSED APPROACH

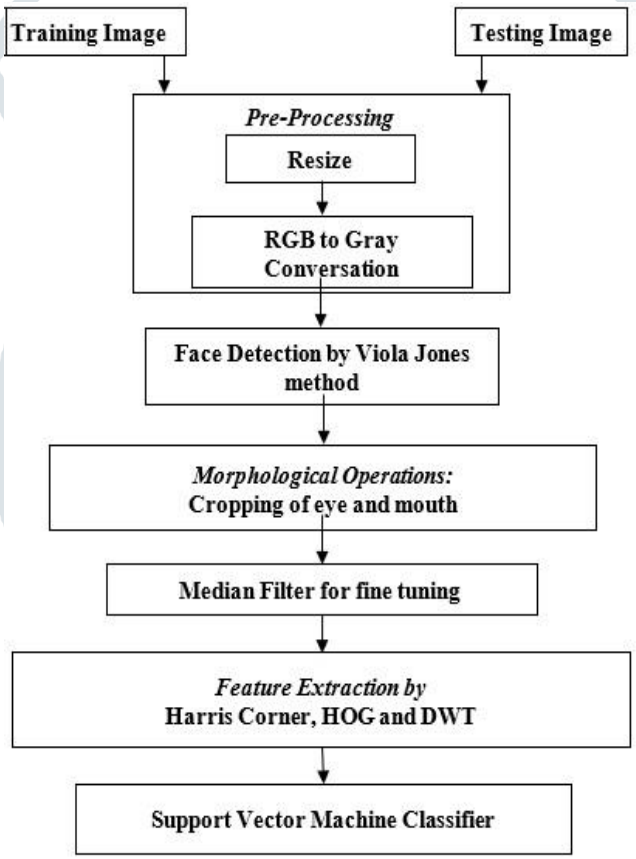

Fig.5: Proposed flow diagram

A. Pre-Processing of Face

1) The size of input photograph is transformed to 224*224 by resize function provided by Matlab Software
2) After completing the resizing process of image. The image is transformed to grayscale by using rgb2gray function of Matlab.

B. Viola Jones Method for Face Detection

The method named Viola Jones which is a real time face detection method used for detection of face and this method was developed in 2001 by Paul Jones. A real time face detection by viola jones based on four fundamental steps.

- Haar Features
- Integral Image
- AdaBoost

- Cascade Classifiers

C. Morphological Operations

Mathematical morphology processes images with the help of previously chosen special shapes, which are generally smaller than the image and are called the structural element [9], which acts as an operator on an image to produce a result. The shape, size and orientation of the structural element are chosen based on prior knowledge taking into account the relevant geometric structures that are present in the image and the objective pursued with the morphological operation implemented [9]. Each structure element requires the definition of a point of origin (or reference) for its application as a morphological operator. This allows the structure element to be related in a particular way to the pixels of the image.

- Dilation
- Erosion

Figure 6 provides an explanation of basic operations of Morphing I.e. Dilation and Erosion:

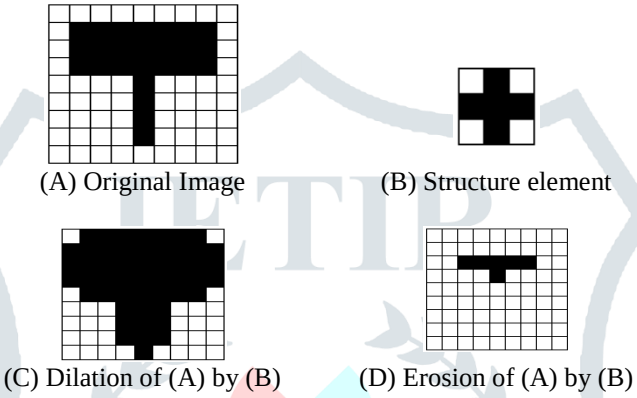

Fig. 6: Demonstration of morphological operations [9]

D. Median Filter

The principle of the Median filter is often defined in the case of a discrete image, whose practical implementation is direct. From the discrete concepts, it is possible to give a continuous version, which will be adapted to the theoretical study of some of its properties [10].

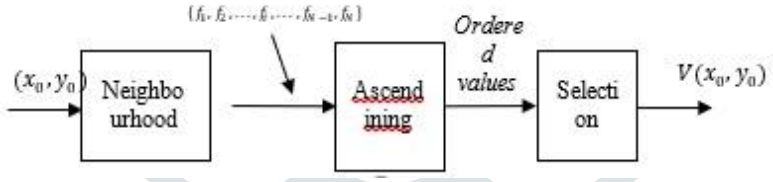

Fig. 7: Median filter

E. Feature Extraction

The goal of extracting feature in biometric recognition field is to express the feature in numerical or symbolic form called encoding. Depending on the case, the values of these features can be real, integer or binary. The vector composed of feature n represents a point in the new space of n dimensions. The steps involved in feature extraction are shown in the flow diagram in Figure 4.1using three feature extraction methods:

- Harris Corner
- DWT (Discrete Wavelet Transform)
- HOG (Histogram Oriented Gradients)

1) Harris Corner Method

Harris corner detection is a post-processing technique used for smile recognition by the lip corner data, eye detection and many more [11]. Harris corner detector has to detect the corners in the iris image I_I.Firstly I_I is used to create gradient image through filtering it with Gaussian Mask. And then the Harris corner method is imposed [11].

2) Discrete Wavelet Transform

The wavelet transform is established on the basis of a signal decomposition by wavelets whose expansion and translation parameters are continuous variables. Wavelets convert the image into a series of wavelets that can be stored more efficiently than pixel blocks. Wavelets have rough edges; they are able to render pictures better.

3) Histogram of Oriented Gradients (HOG)

The oriented gradient [14] is a descriptor using the contour information. For this situation, every window can be portrayed by the local distribution of the introductions and amplitudes. This distribution is defined by means of a histogram. This is modelled by partitioning the location window into an arrangement of cells on which the sufficiency of the gradient for every orientation interval is incorporated over the pixels of the cell.

F. Similarity Measure using Support Vector Machine

Support Vector Machine, or SVMs, are the set of supervised machine learning algorithm designed to solve out regression problems. SVMs are generalization of linear classifiers.

The objective is to find a decision boundary that can separate the cloud of points into two regions by making a minimum of errors i.e. find the optimal hyperplane.

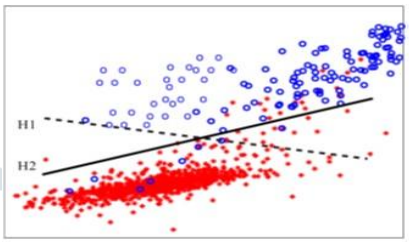

Fig 8: Principle of separating hyperplane

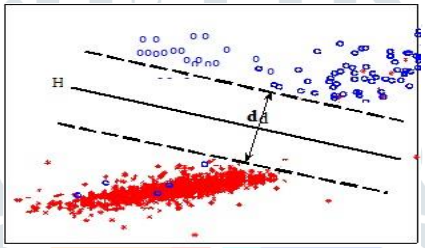

Fig 9: The optimal hyperplane H with the maximum margin d [16]

V. SIMULATION RESULTS

A. Evaluation Parameter used:

		Prediction	
		Positives	**Negatives**
Observation	**Positives**	True Positives (TP)	False Negatives (FN)
	Negatives	False Positives (FP)	True Negatives (TN)

Fig. 10: Confusion matrix example

- TP → number of positives that were correctly classified as positive by the model.
- TN → amount of negatives that were correctly classified as negative by the model.
- FN → number of positives that were incorrectly classified as negative.
- FP → is the number of negatives that were incorrectly classified as positive.

- **Accuracy:** In general, the percentage of the data classified correctly is termed as an accuracy:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

- **Precision (P):** When predicting positives, the percentage classified correctly is termed as precision:

$$P = \frac{TP}{TP+FP}$$

- **Sensitivity:** When the class is positive, the percentage you can classify is referred as sensitivity:

$$S = \frac{TP}{TP+FN}$$

B. Simulation

The simulation is carried out by using MATLAB software image processing toolbox.
Following steps are involved in facial expression recognition. Input image is taken and reduced to 224 × 224 as shown in (Figure 11 and 12).

Fig. 11: Input image

Fig. 12: Image Resized to 224×224

Then using Viola Jones Method face is detected and cropped to small size image which is shown in figure 13 and 14.

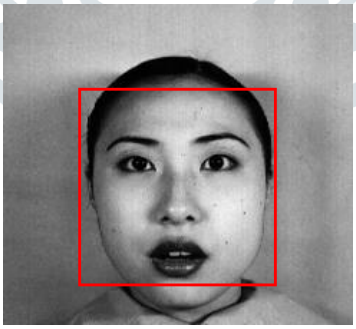

Fig. 13: Detection of face by using Viola Jones method

Fig. 14: Image after cropping

Morphological operations are performed for detecting eyes and mouth, then by using viola jones method eyes and mouth are detected and forming a new cropped image. Feature extraction is performed using Harris Corner, HOG and DWT techniques. This cropped image is again resized to 110×110 for further classification as shown in figure from 15 to 16.

Fig. 15: Detection of Eye and mouth

Fig. 16: Cropped image after eye and mouth detection

Fig. 17: Image Resized to 110×110

The support vector machine based procedure does not use any threshold valve for recognition. The measure or similarity is done by SVM itself. Finally, confusion matrix plot shows the performance of system which is shown below:

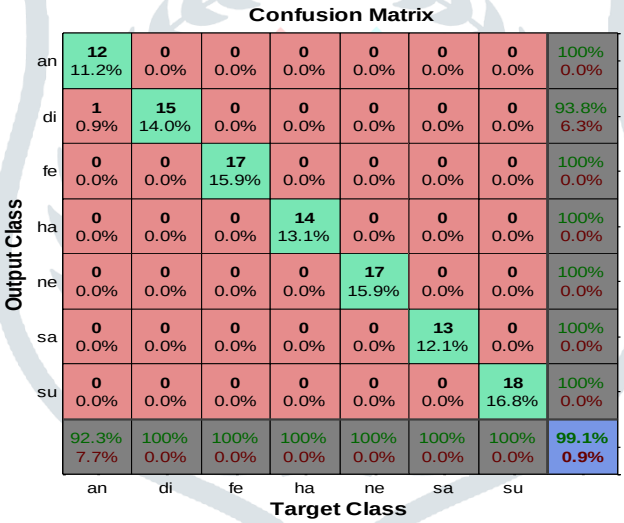

Fig. 18: Confusion matrix for proposed approach for JAFFE database

The accuracy for proposed algorithm is **99.1%**.

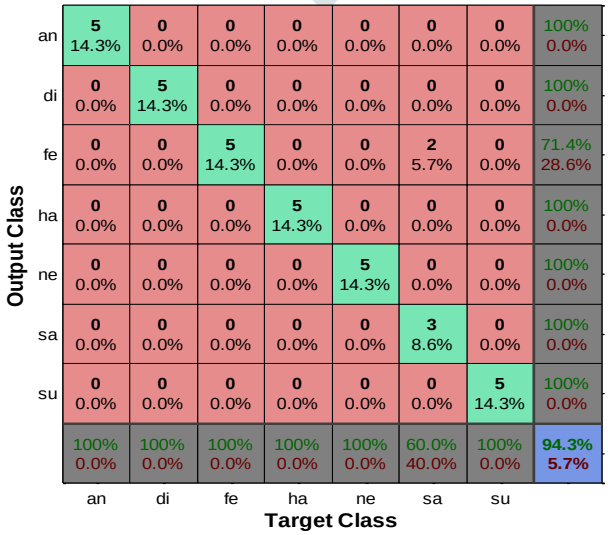

Fig. 19: Confusion matrix for proposed approach for CK database

The accuracy for proposed algorithm is. **94.3%**.

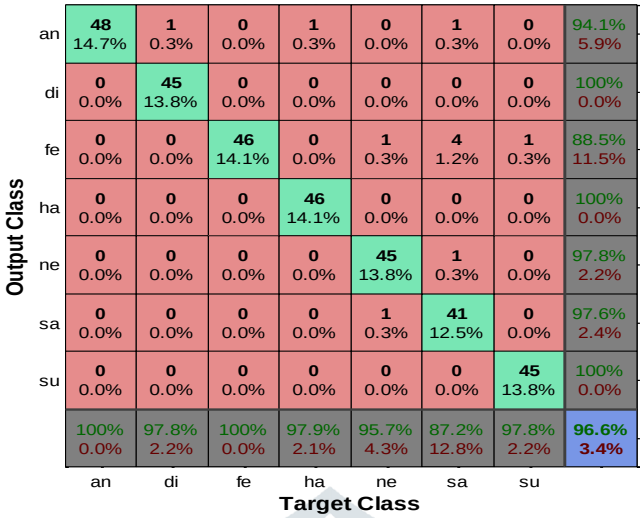

Fig. 20: Confusion matrix for proposed approach for KDEF database

The accuracy for proposed algorithm is. **96.6%**.

VI. Conclusion and Future Scope

This research work represents the facial expression recognition technique using three feature extraction methods Harris Corner, HOG and DWT which is classified by the Support Vector Machine Classifier. Confusion matrix demonstrates that the proposed system gives more accuracy than the previous research works for all three facial expression databases. The maximum accuracy is achieved by this work is 99.1% for JAFFE database and accuracy for CK and KDEF database are 94.3% and 96.6% respectively. Facial expression recognition technique is implemented with few constraints, though the system performs efficiently when input images are frontal and with a particular brilliance. The facial expression recognition algorithms work well in the most practical conditions where the systems are not defined to perform under different physical conditions (for example: night, sunset, dawn). The future systems are expected to identify a human conveniently in distinct operational conditions.

The standard data set available works on particular brightness, contrast, intensity and the techniques available work quite well on the datasets images but can't perform better on real time images so in future we can work on dark/dim/shaded real time images.

The identification systems with robust outputs cannot rest on solo biometric feature because the existence of noise and brilliance may affect the outcome of system. So we can merge it with voice of the person

References

[1] Ioannou, Spiros V., Amaryllis T. Raouzaiou, Vasilis A. Tzouvaras, Theofilos P. Mailis, Kostas C. Karpouzis, and Stefanos D. Kollias. "Emotion recognition through facial expression analysis based on a neurofuzzy network." Neural Networks 18, no. 4 (2005): 423-435.

[2] Fasel, Beat, and JuergenLuettin. "Automatic facial expression analysis: a survey." Pattern recognition 36, no. 1 (2003): 259-275.

[3] Bellamkonda, Sivaiah, and N. P. Gopalan. "Facial Expression Recognition Using Kirsch Edge Detection, LBP and Gabor Wavelets." In 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS), pp. 1457-1461. IEEE, 2018.

[4] Mehta, Neelum, and SangeetaJadhav. "Facial Emotion recognition using Log Gabor filter and PCA." In Computing Communication Control and automation (ICCUBEA), 2016 International Conference on, pp. 1-5. IEEE, 2016.

[5] A. K. Jain, A. Ross and S. Prabhakar, "An introduction to biometric recognition". IEEE Transactions on Circuits and Systems for Video Technology, vol. 14, pp. 4–20, Jan 2004.

[6] Kanade, Takeo, Jeffrey F. Cohn, and Yingli Tian. "Comprehensive database for facial expression analysis." In Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580), pp. 46-53. IEEE, 2000.

[7] Lyons, Michael J., Shigeru Akamatsu, Miyuki Kamachi, Jiro Gyoba, and Julien Budynek. "The Japanese female facial expression (JAFFE) database." In Proceedings of third international conference on automatic face and gesture recognition, pp. 14-16. 1998.

[8] Viola, Paul, and Michael Jones. "Rapid object detection using a boosted cascade of simple features." In Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on, vol. 1, pp. I-I. IEEE, 2001.

[9] Shih, F. "Image processing and mathematical morphology". Boca Raton: CRC press, 2009.

[10] Sezan, M. Ibrahim, and Reginald L. Lagendijk, eds. Motion analysis and image sequence processing. Vol. 220. Springer Science & Business Media, 2012.

[11] Chen, Jie, Li-hui Zou, Juan Zhang, and Li-hua Dou. "The Comparison and Application of Corner Detection Algorithms." Journal of multimedia 4, no. 6, 2009.

[12] Carcagnì, Pierluigi, Marco Del Coco, Marco Leo, and CosimoDistante. "Facial expression recognition and histograms of oriented gradients: a comprehensive study." SpringerPlus 4, no. 1 (2015): 645.

[13] Lajevardi, Seyed Mehdi, and Margaret Lech. "Facial expression recognition using neural networks and log-gabor filters." In Computing: Techniques and Applications, 2008. DICTA'08. Digital Image, pp. 77-83. IEEE, 2008.

[14] Nectarios Rose, "Facial Expression Classification using Gabor and Log-Gabor Filters", IEEE 7th International Conference on Automatic Face and Gesture Recognition, 2006.

[15] Sanchez-Mendoza, David, David Masip, and AgataLapedriza. "Emotion recognition from mid-level features." Pattern Recognition Letters 67, pp. 66-74, 2015.

[16] Li, Li. "Support vector machines." In Selected applications of convex optimization, pp. 17-52. Springer, Berlin, Heidelberg, 2015.