Use of KNN Classifier for Emotion Recognition Based on Distance Measures

Vishal D. Bharate, Devendra S. Chaudhari, Mayur D. Chaudhari

Abstract**:** ***Human-computer interaction (HCI), in recent times, is gaining a lot of significance. The systems based on HCI have been designed for recognizing different facial expressions. The application areas for face recognition include robotics, safety, and surveillance system. The emotions so captured aid in predicting future actions in addition to providing valuable information. Fear, neutral, sad, surprise, happy are the categories of primary emotions. From the database of still images, certain features can be obtained using Gabor Filter (GF) and Histogram of Oriented Gradient (HOG). These two techniques are being used while extracting features for getting the expressions from the face. This paper focuses on the customized classification of GF and HOG using the KNN classifier.GF provides texture features whereas HOG finds applications for images exhibiting differing lighting conditions. Simplicity and linearity of KNN classifier appeals for its use in the present application. The paper also elaborates various distances used in KNN classifiers like city-block, Euclidean and correlation distance. This paper uses Matlab implementation of GF, HOG and KNN for extracting the required features and classification, respectively. Results exhibit that the accuracy of city- block distance is more .***

Keywords* : *Gabor Filter, Histogram of Oriented Gradient, Human-computer interaction, k-nearest neighbors.

I. INTRODUCTION

One of the recent topics in today's research is the automated facial expression recognition. Many challenges are being faced by researchers in the field of the human facial recognition system. They need to be addressed with the incorporation of optimized and robust techniques which include a choice of appropriate classifier and application of feature engineering. The automation feature of machine learning algorithms makes them an efficient and effective choice for the detection of human emotions. There are three phases in emotion recognition based on the facial expressions of a human being. In the first phase, the still images from the database are captured and processed. During the second phase, the facial expressions are processed for extracting features with the use of information from the edges of images and other facial data. The result of the second phase are the feature vectors containing information about various human emotions extracted from expressions of the face. Feature vectors for classification of emotions from facial expressions are used in the third phase.

There are various constituent parts in human emotions based on human physiology. During social interactions, these constituent parts include 7% verbal, 38% vocal and 55% facial part [6, 10]. Facial expressions[9] being a dominant part of expressing emotions, it has been considered for various research works. Various researchers have emphasized the use of geometric features for emotion recognition based on facial expressions[15]. This approach has exhibited less efficiency for emotion recognition [4]. Appropriate features need to be selected for efficient emotion recognition based on a specific application. The various challenges faced by researchers during emotion recognition based on facial expressions are face shape, size, viewing angle, change in object pose, exposure to illumination[7,11,12] and occlusion. In this paper, an increase in the efficiency of facial expression detection in HCI using machine learning is attempted using HOG and GF. The paper explains the use of KNN classifier that applies various distances like city-block, Euclidean and correlation. Amongst these distances city-block distance provides robust and significant results during the classification of human emotions.

The paper is organized into the following parts: Section II describes current practices and Section III presents the framework of proposed work, along with experimental results.

II. RELATED WORK AND PROPOSED FRAMEWORK

This section presents the proposed methodology of HCI. The section begins with the block schematic of implemented work, followed by various steps involved in feature engineering and classification of emotions based on facial expressions.

The block schematic of the implemented framework of HCI for human emotion detection has been presented in Fig. 1. Preprocessing, feature engineering (extraction) and classification[8] form the major components of the proposed framework. The proposed framework has been further explored in detail in the subsequent sections.

A. Pre-Processing

In this, local normalization reduces the noise, and the degraded image component so that the system would give better performance in emotion recognition [5].

Revised Manuscript Received on October 15, 2019

Vishal D. Bharate, Department of Electronics and Telecommunication, Government College of Engineering, Amravati/ Sinhgad Academy of Engineering, Pune, India. Email: vishalbharate@gmail.com

Devendra S. Chaudhari, Department of Electronics and Telecommunication, Government College of Engineering, Jalgaon, Jalgaon, India. Email:ddsscc@yahoo.com

Mayur D. Chaudhari, Data Architect, Parkar Labs, Pune, India. Email: chaumayu@gmail.com

Local mean along with a variance of an image under consideration have been normalized using the local normalization algorithm.

$$n(a,b) == i(a,b) - mi(a,b)/\sigma i(a,b) \quad (1)$$

Normalized form of an image under consideration has been presented in (1). Here a and b map to x and y-axis respectively on the Cartesian coordinate system. In (1), $i(a,b)$ contains pixel information of the image under consideration, $mi(a,b)$ and $\sigma i(a,b)$ represent respectively the estimations of the local mean and local variance of $i(a,b)$.

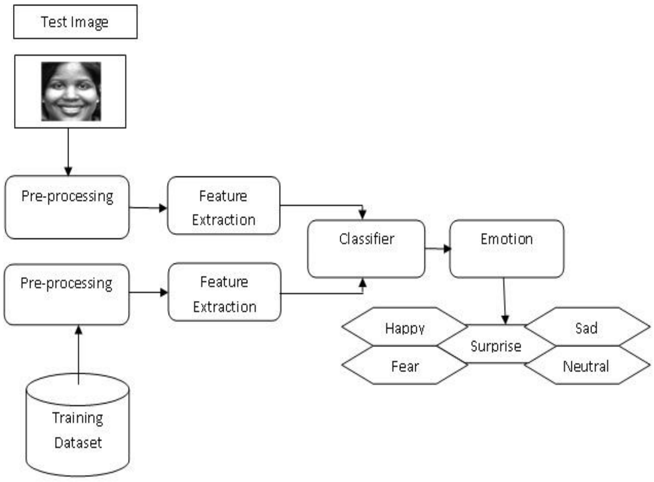

Fig. 1.Block Schematic of implemented HCI

B. Feature Extraction

The original image has been pre-processed, and further various features have been extracted using HOG and GF. The following subsection provides a brief overview of HOG and GF used during feature extraction.

1) Histogram of Oriented Gradient

One of the feature descriptors used during feature extraction is HOG[1,8]. HOG has been extensively employed in computer vision and image processing applications. The occurrence of gradient orientation has been counted in a localized region. The localized intensity distribution has been computed by using HOG.

The pre-processed image forming the input to the feature extraction block has been divided into a number of small cells. The pixels of every cell has been further compiled using histogram gradient [1]. HOG feature descriptor is formed by concatenating the histograms gradients so compiled. Amongst the three transformations, namely, object orientation, geometric and photometric transformation, which operate on local cells, the HOG feature descriptor is invariant to geometric and photometric transformations.

Various steps in the implementation of the HOG are described in the following text:

Step 1: Gradient computation

The X and Y gradient values have been computed initially. Using Sigma as 0.5, the Gaussian filter has been applied for smoothing of the image under consideration. Using the magnitude and angle of every edge, the edge gradients have been computed.

Step 2: Orientation binning

This step involves the computation of histogram for each cell after gradient computation. Weighted votes are computed for each pixel in every cell by applying the gradient computation step, which contributes towards the formation of histogram channels. The gradient magnitude provides weighted votes. The shape of the cell, either radial or rectangular and gradient of the cell, either 'signed' or 'unsigned', decide the even spreading of histogram channels between either 0-180 or 0-360.

Step 3: Formation of descriptor blocks

The contrast and illumination of the image are observed by locally normalizing the oriented gradient values. To aid this process, the spatial blocks are formed by combining the pixels of each cell. The histogram gradients from each normalized cell are concatenated to form histogram descriptor.

Step 4: Block normalization

The condition for the normalization process is given in (2).

$$Lnorm \rightarrow f = v_{nn} / \|v_{nn}\| + e \quad (2)$$

where,

v_{nn} is the non-normalized vector and

e is the error vector.

2) Gabor Filter

The edge detection of an image is performed by extracting features with the aid of a Gabor filter [2]. The Gabor filter's impulse response is computed using the sinusoidal wave and the Gaussian function. The convolution function is applied to Fourier transforms of Gaussian and harmonic functions. This results in the Fourier transform of the impulse response of the Gabor filter. Equation (3) represents the real and imaginary components of the Gabor filter.

$$g(p,q,\lambda,\theta,\phi,\gamma) = e^{(p^2+q^2\lambda^2/2\sigma^2)} + e^{(i(2\pi p/\lambda+\psi))} \quad (3)$$

where,

λ = Wavelength

θ = Orthogonal to parallel stripes orientation

ϕ = Offset in the phase

γ = Spatial aspect ratio and

σ = Standard deviation

Here p and q map to x and y-axis respectively on the Cartesian coordinates system. Equation (3) is useful in computation of orientation of orthogonal to the parallel stripes of the Gabor function and the wavelength of the sinusoidal factor. The offset in the phase is computed using the standard deviation of the Gaussian envelope. Spatial aspect ratio, that specifies Gabor function, has been used in creating the Gaussian envelope. The features from any image are extracted using a Gabor filter that has different orientation and frequencies.

The Gabor filter is represented by using following equations:

$$G1[i,j] = Ae^{((p^2+q^2)/2\sigma^2)} \cos(2\pi f(i\cos\theta + j\sin\theta)) \quad (4)$$

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

$$G2[i, j] = Be^{-((p^2+q^2)/2\sigma^2)} \cos(2\pi f(i\cos\theta + j\sin\theta)) \quad (5)$$

where,

A and B= normalizing factors

f =frequency

θ = Normal to parallel stripes orientation and

σ = Standard deviation

All submitted paper should be cutting edge, result oriented, original paper and under the scope of the journal that should belong to the engineering and technology area. In the paper title, there should not be word 'Overview/brief/ Introduction, Review, Case study/ Study, Survey, Approach, Comparative, Analysis, Comparative Investigation, Investigation'.

C. Classification

The primary purpose of classification in emotion recognition is to separate out various emotion classes. The classification can be performed on binary or multiple classes. Identification of classes is the primary step in classification. This paper presents KNN classifier being a simple classifier based on the criteria of the neighborhood.

KNN Classifier

In pattern recognition, K-Nearest Neighbor [5, 13, 14] classifier applies a non-parametric method for classification of various classes. The classifier checks for the classes that are nearer to the test pattern. The classifier input is formed by using the K closest set of training patterns. The classification algorithm functions based on the odd number of training patterns that form the nearest neighborhood. The output of the classifier has been computed using class of the maximum number of training patterns that form the nearest neighborhood. The classification of feature vector into a particular class is hence performed by comparing it with an odd number of nearest neighboring training feature vectors. For example, for K=1, it is a single nearest neighbor, while for K=3, the classifier compares test vector with three training feature vectors and finds a maximum match of testing feature vector with other three training feature vectors, and for K=5,7,9 and so on. The entire classification algorithm has been presented in Fig. 2.

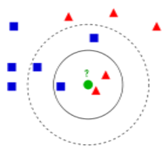

Fig. 2.KNN classification[18]

In Fig. 2, three geometric patterns, namely triangle, rectangle and circle, are presented as feature vectors. The geometric patterns are chosen so as to provide ease of understanding. The training features are represented by a triangle and rectangle, whereas the circle forms the testing feature. As shown in Fig. 2, the test feature (circle) is compared with three nearest neighbors, two out of which are triangles and one is a rectangle. According to the KNN classification algorithm, the circular pattern is classified into triangle class, as two out of the three nearest neighbors are triangles. The methodology presented applies to all the real-time classes.

The KNN classifier uses the following distance metrics [16]:

1. City Block Distance

In this metric, the absolute value of the difference between Cartesian coordinates is summed up over a limited range to get the neighborhood distance value.

$$d_{st} = \sum_{j=1}^{n} \left| p_{sj} - q_{tj} \right| \quad (6)$$

where,

p and q= Cartesian coordinates

The city block distance is computed using (6).

2. Euclidean Distance

This distance metric is computed as the multiplication of the differences between Cartesian coordinate of one point and the transpose of the Cartesian coordinate of the same point.

$$d_{st}^2 = (p_s - q_t)(p_s - q_t)' \quad (7)$$

where,

p and q= Cartesian coordinates

Euclidean distance is computed using (7).

3. Correlation Distance

This metric is computed using the correlation of statistical dependence between two vectors.

$$d_{st} = \left(1 - \frac{(p_s - \bar{p}_s)(q_t - \bar{q}_t)'}{\sqrt{(p_s - \bar{p}_s)(p_s - \bar{p}_s)'}\sqrt{(q_t - \bar{q}_t)(q_t - \bar{q}_t)'}} \right) \quad (8)$$

where,

$$\bar{p}_s = \frac{1}{n}\sum_j p_{sj} \text{ and}$$

$$\bar{q}_t = \frac{1}{n}\sum_j q_{tj}$$

Correlation distance is computed using (8).

In the research area of investigation of pattern recognition, the efficiency of a classifier is judged by various measures like confusion matrix, classification accuracy and time required for classification of a sample test image. Confusion matrix provides the percentage of misclassification. Classification accuracy is computed using (9).

$$Accuracy = \frac{P_T + N_T}{P_T + N_T + P_F + N_F} \quad (9)$$

where,

P_T = Positively True

N_T = Negatively True

P_F = Positively False

N_F = Negatively False.

Every classification algorithm requires a finite time to classify the test image. This time provides the speed of classification of the algorithm in the testing phase. For a classification algorithm to work efficiently, this time should be as short as possible.

III. RESULT AND DISCUSSION

Results after experimentation over proposed HCI system have been discussed in this section. The section also describes the database and system specifications where experimentation is carried out.

A. Database and Experimental setup

The standard dataset JAFEE database [17] comprising of 213 static images has been used for experimentation purpose. The database comprises of images in TIFF format having dimensions 256×256 pixels and storage size in the range between 60 to 70 KB. The use of standard database facilitates accuracy, more specific results and comparison of results. The employed database is classified into testing and training datasets. Training dataset contains 140 images, whereas the testing dataset uses the remaining 73 images. The experimental setup consists of Matlab version 2014a software installed on a laptop having Windows 7 OS, i3 processor and 3 GB RAM. The purpose of Matlab software is to run the simulation of the proposed algorithm on the said dataset.

B. Results and Confusion Matrix

In this work, KNN classifier has been applied to training and testing datasets. In the training phase, the HOG feature vectors of training dataset are computed. The testing phase involves computation of the HOG features of the image under test and comparing with the HOG features of all the images in the training dataset. The best match is based on the closest distance of the testing feature from that of training dataset feature. Using a similar methodology, GF features are computed and applied to training and testing datasets. KNN classifier is applied on GF feature vectors, and the results are presented in terms of the confusion matrix shown in Table I. It can be seen from Table I that Happy and Surprise class recognition rate is maximum. Another finding from Table I is that sad and fear expressions are not distinguished accurately due to the high level of similarity in the feature vectors.

Table- I: Confusion matrix for KNN applied to GF features

Emotion	Fear	Neutral	Sad	Surprise	Happy
Fear	70%	0	20%	10%	0
Neutral	0	60%	20%	0	20%
Sad	40%	0	60%	0	0
Surprise	10%	10%	0	80%	0
Happy	0	10%	10%	0	80%

On similar lines, KNN classifier is applied on HOG feature vectors, and the results are presented in Table II, in terms of the confusion matrix.

Table- II: Confusion matrix for KNN applied to HOG features

Emotion	Fear	Neutral	Sad	Surprise	Happy
Fear	80%	0	20%	0	0
Neutral	0	70%	10%	20%	0
Sad	20%	0	60%	0	20%
Surprise	0	0	0	100%	0
Happy	0	10%	0	0	90%

From Table II, it is seen that for Happy and Surprise classes recognition rate has increased in comparison with that of GF feature vector. From the confusion matrix presented in Table II, it can be observed that sad and fear emotions are not distinguishable due to the high level of similarity of feature vectors.

Further in this work, various distance metrics namely city-block, Euclidean and correlation are computed for KNN classifier. The accuracies of the results of the distance computation of the said distance metrics are presented in Table III.

Table- III: KNN classification accuracy based on various distance metrics.

Distance / Feature	Correlation Distance	Euclidean Distance	City-Block Distance
Gabor Filter	68%	69%	80%
HOG Transform	69%	65%	82%

It can be seen from Table III that the accuracy of HOG transform is much more for city-block distance as compared to that for than that for Euclidean and correlation. Amongst HOG and GF transforms, excepting the case of Euclidean distance, the accuracy of HOG transform is more than that for GF transform.

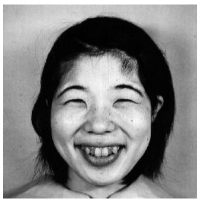

Fig. 3. Test image[17] for computation of execution time

Fig. 3 depicts a sample image taken from test dataset for computation of execution time[3] of the KNN classifier. Using the HCI presented in this paper, this image is correctly classified into happy class. The image is processed with the help of KNN classification algorithm presented in this work, and the execution time is computed using Matlab. The execution time is computed for both HOG and GF features using KNN classifier and using Table IV, the results are presented.

Table- IV: Execution time for GF and HOG features computed using KNN classifier

Feature	Execution Time
Gabor Filter	2.6064
HOG Transform	0.7684

From Table IV, it can be seen that the execution time of the HOG transform is much less than that for GF transform for the selected sample image. After carrying out experimentation on all the test images from the database, it has been observed that 72% of the test images exhibit execution time similar to that depicted in Table IV. Hence, the HOG transform is much faster than GF transform in case of KNN classifier.

IV. CONCLUSION

In this paper emotions are recognized from still images of face using JAFEE database. GF and HOG are used for extracting various features. KNN classifier is used for identifying individual emotion class from features extracted using GF and HOG.

The HCI system presented in this paper is able to detect five emotions, namely fear, neutral, sad, surprise and happy, based on facial expressions. From the simulation results, it can be seen that happy and surprise classes are well distinguished due to the non-similarity in facial expressions. On the other hand, sad and fear classes are not distinguished due to the high level of similarity in the facial expressions. The accuracy of HOG is 10-15% more as compared to that of GF. Although the combination of HOG and KNN exhibits good accuracy, some other combinations like fusion of features along with various classifiers may provide better results. The execution time for HOG is less than GF. In terms of distance metrics, HOG proves better for city block distance than Euclidean or correlation distance.

REFERENCES

1. O. Dniz, G. Bueno, J. Salido and F. Torre, "Face recognition using Histograms of Oriented Gradients," Pattern Recognition Letters, vol. 32, no. 12, pp. 1598-1603, Sep. 2011.
2. R. Singh, S. Jadhav, S. Sharma and S. Gupta, "Average Gabor Wavelet Filter Feature Extraction Technique for Facial Expression Recognition," International Journal of Computer Science and Network (IJCSNS), vol.15, no. 5, May. 2015.
3. J. Nagi, S. Ahmed and F. Nagi, "A MATLAB based Face Recognition System using Image Processing and Neural Networks," International Colloquium on Signal Processing and its Applications, 2008.
4. R. Calvo and S. D'Mello, "Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications," IEEE Transactions on Affective Computing, vol. 1, no. 1, pp. 18-37, Jan. 2010.
5. G. Liu, X. Zhang and S. Zhou, "Multi-class Classification of Support Vector Machines Based on Double Binary Tree," Fourth International Conference on Natural Computation, Jinan, 2008, pp. 102-105.
6. J. Vaibhavkumar, M. Mistry and M. Goyani, "A literature survey on Facial Expression Recognition using Global Features" International Journal of Engineering and Advanced Technology (IJEAT), vol. 2, no. 4, Apr. 2013.
7. S. Sagarika and P. Maben, "Face Recognition and Facial Expression Identification using PCA", IEEE International Advance Computing Conference (IACC), Gurgaon, India, 2014.
8. A. Bhadu, R. Tokas and V. Kumar, "Facial Expression Recognition Using DCT, Gabor and Wavelet Feature Extraction Techniques," International Journal of Engineering and Innovative Technology, vol. 2, no. 1, Jul. 2012.
9. K. Ko and K. Sim, "Development of a Facial Emotion Recognition Method Based on Combining AAM with DBN," International Conference on Cyberworlds, Singapore, 2010, pp. 87-91.
10. S. Pande and S. Shinde, "A Survey on Emotion Recognition with respect to Database and Various Recognition Techniques", International Journal of Computer Applications, vol. 58, no. 3, 2012.
11. C. Sumathi, T. Santhanam and M. Mahadevi, "An Automatic Facial Expression Analysis A survey," International Journal of Computer Science & Engineering Survey (IJCSES), vol. 3, no. 6, Dec. 2012.
12. Y. Tie and L. Guan, "Automatic face detection in video seq. using local nor. & OAC tech," Journal of Pattern Recognition, vol. 42, no. 9, pp. 1859-1868, 2009.
13. A. Smeulders, M. Worring, S. Santini, A. Gupta and R. Jain, "Content-based image retrieval at the end of the early years," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 12, pp. 1349-1380, Dec. 2000.
14. Z. Guo, L. Zhang and D. Zhang, "A Completed Modeling of Local Binary Pattern Operator for Texture Classification," IEEE Transactions on Image Processing, vol. 19, no. 6, pp. 1657-1663, Jun. 2010.
15. Z. Guo, L. Zhang and D. Zhang, "Rotation invariant texture classification using LBP variance (LBPV) with global matching," Pattern Recognition, vol. 43, no. 3, pp. 706-719, Mar. 2010.
16. K. Chomboon, P. Chujai, P. Teerarassamee, K. Kerdprasop and N. Kerdprasop, "An Empirical Study of Distance Metrics for k-Nearest Neighbor Algorithm," International Conference on Industrial Application Engineering, 2015
17. M. Kamachi, M. Lyons and J. Gyoba, "The Japanese female facial expression (JAFEE) database," International Conference on Automatic Face and Gesture Recognition, 1997.
18. Srivastava, T. (2018, March 26). Introduction to k-Nearest Neighbors: A powerful Machine Learning Algorithm (with implementation in Python & R). Retrieved August 10, 2019, from analyticsvidhya.com: https://www.analyticsvidhya.com/blog/2018/03/introduction-k-neighbours-algorithm-clustering/

AUTHORS PROFILE

Vishal D. Bharate obtained BE, from Shivaji University, Kolhapur and M.Tech. from Dr. Babasaheb Ambedkar University, Lonere. He is pursuing Ph. D. from Government College of Engineering, Amravati. He has been engaged in teaching, research for about 12 years. Presently he is working as Assistant Professor, Department of Electronics and Telecommunication Engineering at Sinhgad Academy of Engineering, Pune, Maharashtra, India. He has published research papers and presented papers in international conferences at Chennai, Kanyakumari. He has published 3 Journal and 5 Conference papers. His present research and teaching interests are in the field of Digital Image Processing, Digital Signal Processing, Computer Vision and Machine Learning.

Devendra S. Chaudhari obtained BE, ME, from Marathwada University, Aurangabad and PhD from Indian Institute of Technology, Bombay, Powai, Mumbai. He has been engaged in teaching, research for about 30 years. He has worked as Head Electronics and Telecommunication, Instrumentation, Electrical, Research and incharge Principal at Government Engineering Colleges. Presently he is working as Head, Department of Electronics and Telecommunication Engineering at Government College of Engineering, Jalgaon. Dr. Chaudhari published research papers and presented papers in international conferences abroad at Seattle, USA and Austria, Europe. He is fellow of IE, IETE and life member of ISTE, BMESI and member of IEEE (2007). He is the recipient of Best Engineering College Teacher Award of ISTE, New Delhi, Gold Medal Award of IETE, New Delhi, Engineering Achievement Award of IE (I), Nashik. He has organized various Continuing Education Programmes and delivered Expert Lectures on research at different places. He has also worked as ISTE Visiting Professor and visiting faculty member at Asian Institute of Technology, Bangkok, Thailand. His present research and teaching interests are in the field of Biomedical Engineering, Digital Signal Processing and Analogue Integrated Circuits.

Mayur D. Chaudhari received the B.E. degree in computer engineering from the Government College of Engineering, Jalgaon, Maharashtra, India, in 2016, and M.E. degree in computer engineering from the Pune Institute of Computer Technology, Pune, Maharashtra, India, in 2018. He has been engaged in research for about 4 years. He joined the Data Science Company named Chistats, as a Data Scientist in 2018. Currently, he is working as a Data Architect with Parkar Labs. He has 3 international level publications in international conference and journal. His main areas of interest are Machine Learning, Image Processing, Deep Learning, Natural Language Processing and Text Processing.