A Review on Facial Emotion Recognition Techniques

Jacintha .V, Judy Simon, Tamilarasu .S, Thamizhmani .R, Thanga yogesh. K and J. Nagarajan

***Abstract*—As a major phase skin Recognition, together with elliptical boundary model, is accomplished. Further, facial Feature Identification process is carried out. The next step is to initiate a technique for extorting geometric and anthropometric facial characteristics. At last we train as well as test the classifiers. We accomplished a categorization precision of 58.6% for six types of emotions (bliss, anguish, curiosity, despair, fury, hatred) and mean efficiency of 96.8% for two emotions (bliss and curiosity).The current study utilizes interest points as markers in face images that are damaged by few emotions as well as correlates its location to that of a normal expression. The outputs are viewed in contrast with Paul Ekman's FACS (Facial Action Coding System) tool to check on the efficacy of the algorithm .The automated identification of face expressions utilizing image template matching method faces various issues pertaining to facial features and recording circumstances. Although this field has reached great heights , choosing of features as well as categorization method for emotion identification , till today remains an unsolved mystery. To suppress feature outliers, the proposed technique comprises of pixel normalization which is used to eliminate intensity offsets backed up using a Min-Max metric in a nearest neighbor classifier. The proposed Min-Max classification technique has an efficiency of 92.85% to 98.57% when checked on JAFFE database. Classification task also done using KNN, SVM and Bagged Tree Classifier.**

***Index Terms*—Emotion recognition, FACS,JAFFE, Min-Max classifier, KNN, SVM.**

I. Introduction

MUSIC seems to be a high priority for majority of the people. Latest technologies allows personal gadgets (e.g., smart phone, tablet, etc) that possess an Internet connection to approach billions of songs simultaneously. As a part of emotion, music generally signals our choice of music in which a content-based approach automatically indexes and systematizes music media. In the digital era, emotion plays a major role for enabling recovery of music. With recent researches in the field of music information retrieval, we found a new possibility that music can be automatically analyzed and understandable by the computer to some extent. Emotion recognition serves as a core element in human-computer interaction, which makes the task quiet natural [1].

Jacintha .V and Judy Simon are with Assistant Professor, Department of Electronics and Communication Engineering, Jeppiaar Maamallan Engineering College, Chennai.{email : jacinthaecegmail.com}.

Tamilarasu . S, Thamizhmani . R, Thanga yogesh. K and J. Nagarajan are with Department of Electronics and Communication Engineering, Jeppiaar Maamallan Engineering College, Chennai, India. {email : ktyyogesh@gmail.com}

Human emotions can be brought out in several forms , through way of talk, expressions of the face, gestures, also for instance, waving of the hand to bid goodbye to a friend[1]. The facial emotion recognition has enormous uses in varid fields, like healthcare, robotics, computer vision, vigilance systems, machine learning, artificial intelligence, communication, psychology, smart environment, safety and embedded systems. In emotion recognition we normally analyze the physiological phase of an individual, i.e. variations in body heat, skin resistance (GSR), modifications in either electrocardiographic signal (ECG) , blood pressure . These signals can be effortlessly viewed when two people interact with each other – the voice of the individual and his/her facial image [2]. This study, started few decades back, permits to detect six elementary emotions, which are bliss, anguish, curiosity, despair, fury, hatred. The rest of this paper Existing System is briefly explain in Section II and concludes the paper in Section III. Finally future work is discussed in Section IV.

II. Existing System

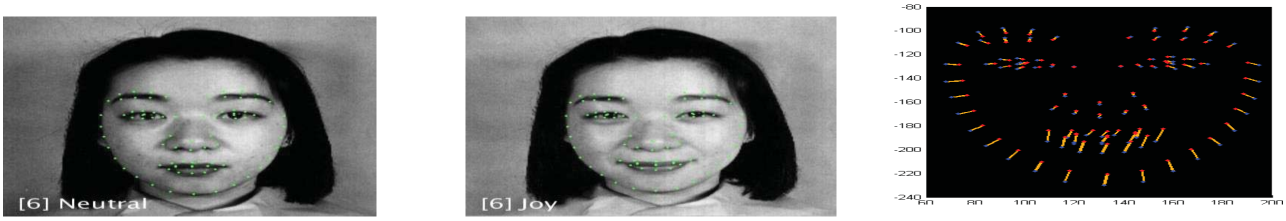
A. A technique for Facial Emotion Identification depending on Interest points

A face image is initially examined using a landmark detector and a set of interest points is fetched. These interest points are correlated with the interest points denoting an individuals neutral expression[3]. Deviations in interest points location (Δx and Δy) are later examined using a classifier to certainly find out what emotion is expressed.

We have chosen six people from this world, who are having seven varid expressions (among which one is taken to be neutral).Images are later treated utilizing , Application Programming Interface (API) Face++, a CNN based technique of cognitive services in the cloud [4]. Face images are given as input to the Face++ servers, that outputs the x and y coordinates of eighty three interest points for every face that is identified. Within each facial expression shows in Fig. 1 there are seven key expressions: bliss, anguish, curiosity, despair, fury, hatred, and contempt[5]. Next is verification of the relationship between variations in the interest points rendered by Face ++ and the categorization done by individuals of the expressions stored in the server. For every interest points, the variation within its location when there is no expression on the face , that is when the face is normal and its location in each of the balance six expressions was computed (Fig. 2(a) and (b). Hence, these variations, taken as an average amidst the six individuals, were plotted.

978-1-5386-7595-3/19/$31.00 ©2019 IEEE

Fig. 1. Facial emotion realization process



(a) (b)

Fig. 2. (a)Illustration of interests points position in neutral and joy expressions (b)Vary vectors amidst interest points.

B. Results

Bliss: At the points 74 and 28 we anticipate a meager raise of the lower eyelid and at points 47 and 38 an enhancement at the edge of the lips. The outputs match with the anticipations.

Anguish: At the points 24 and 66, we find the internal lifting muscles of the eyebrows get triggered, though they should have gone up, they tend to droop lightly, mostly by the movement of the nearby muscle group. The edges of the mouth (points 47 and 38) seem to go up lightly, when they were supposed to go down.

Curiosity: Points 24 and 66 properly rise, which is due to the effect of internal elevators of the eyebrows. The elevator muscles present in the upper eyelid in Fig. 3 and Fig. 4 (points 35 and 81) undergo a raise, much deeply towards the right side when compared to the left side.

Despair: which is the most difficult expression to identify because many set of muscles work at the same time, which rarely cancels out the observable effects within themselves.

Fury: At the points 25, 26 and 72 the eyebrows are lowered down. The upper eyelids shrink and sharpen at points 35 and 81, as predicted.

Dislike: The transversal muscle of the nose enables points 63 and 64 to enhance as anticipated whereas points 38 and 47 near the edges of the lips seem to enhance when they must have lowered Fig. 5.

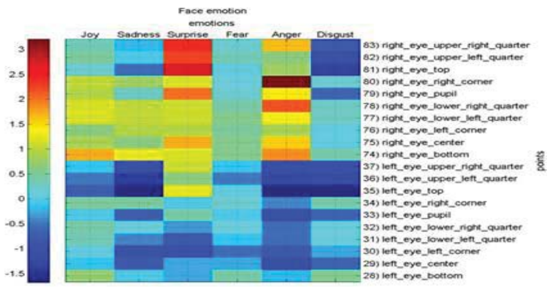


Fig. 3. Δy of the interest points relating to the eye

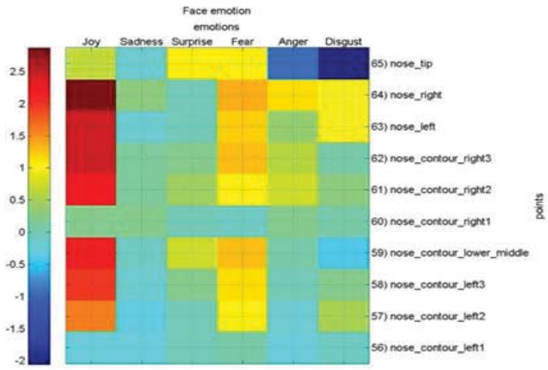


Fig. 4. Δy of the points relating to the nose

C. Facial Emotion Recognition utilizing MIN-MAX Similarity Classifier

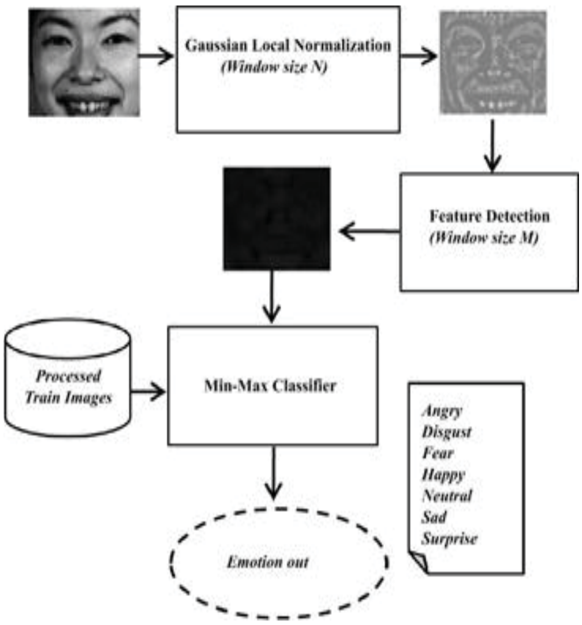


Fig. 5. Overview of the Novel Emotion realization system

The key task of the proposed technique is by utilizing a Min-Max similarity Nearest Neighbor classifier to excerpt the spatial variation of normalized pixels present in the face image and to identify which class it belongs to Experiments are carried out with the aid of images from the JAFFE database in Fig. 6(a), (b) and (c). The JAFFE [6] database consists of 6 elementary facial expressions and neutral faces of 10 women with a total of 213 images. The database consists of the original image of size 256 × 256 pixels , in order to retain only the required data from the face area, the image is cropped to a size of 102 × 115 pixels[7-8].

1. Pre-processing

Fig. 6. (a)An image from JAFFE server having various lighting situations (b) Normalized images of above example images fetched by implementing Gaussian normalization utilizing local mean and local standard deviation obtained using a window of size N=11. (c) Feature detected images from the normalized image by computing standard deviation utilizing a window of size M=11.

The inter-class feature mismatch is brought in by illumination changes , that results in imperfections in identification of emotion segregating features from the facial images. Hence, image normalization is crucial to minimize the inter-class feature mismatch which could be seen as intensity offsets. Inside a local region the intensity offsets are said to be standard and hence Gaussian normalization utilizing mean and standard deviation are carried out.

2. Feature detection

The key features that are helpful in facial emotion identification are eyes, eyebrows, cheeks and mouth areas. Feature detection is carried out by computing the standard deviation of normalized image aided with a window of size M×M.

3. Emotion Categorization

For the recognition phase, we introduce a Min-Max correlation metric in a Nearest Neighbor classifier scheme. The key principle utilized in this method is that the proportion of the minimum variation to the maximum variation of 2 pixels would render a unity result for same pixels and if the pixels are different the output is less than unity[9].

$$s(i,j) = \left[\frac{\min[train(i,j), test(1,j)]}{\max[train(i,j), test(1,j)]}\right]^{\alpha} \tag{1}$$

$$z(i) = \sum_{j=1}^{m \times n} s(i,j) \tag{2}$$

$$out = \max(z(i)) \tag{3}$$

4. Algorithm 1 Emotion Recognition utilizing Min-Max classifier

Require: Test image Y,Trained images Xt,trainlen,window size

N and M

1: Crop the images to a dimension of m×n

2: For t=1 to trainlen do

3: $C(i,j) = \frac{X(i,j)-\mu(i,j)}{6\sigma(i,j)}$

4: $W(i,j) = \sqrt{\frac{1}{M^2}\sum_{k=-b}^{b}\sum_{h=-b}^{b}[C(k+i,h+j)-\mu(i,j)]}$

5: Save the value of W to a sequence train of dimension trainlen×m×n.

6: end for

7: for t=1 to trainlen do

8: $V(i,j) = \frac{Y(i,j)-\mu(i,j)}{6\sigma(i,j)}$

9: $test(i,j) = \sqrt{\frac{1}{M^2}\sum_{k=-b}^{b}\sum_{h=-b}^{b}[V(k+i,h+j)-\mu(i,j)]}$

10: $s(t,j) = [\frac{\min[train(t,j), test(1,j)]}{\max[train(t,j), test(1,j)]}]^3$

11: $z(t) = \sum_{j=1}^{m\times n} s(t,j)$

12: $out = \max(z(t))$

13: end for

5. Results

To check the Efficiency of the Novel Min-Max classifier , we are going to correlate shows in Fig. 7 it with various classifiers like Nearest Neighbor and Random Forest. Table I displays the fetched accuracy percentage.

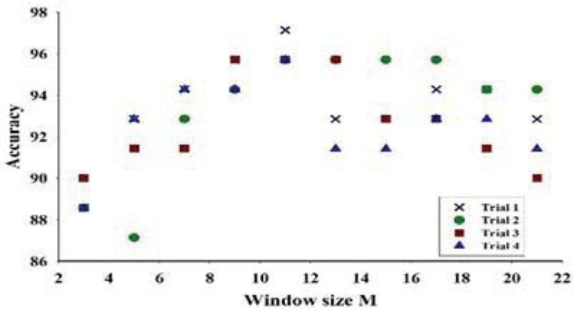


Fig. 7. The accuracy rates fetched for various values of feature detection window size M varying from 3 to 21.

TABLE I
VALIDATION OF FEATURE DETECTED IMAGES ON DIFFERENT CLASSIFIERS

Classifier	Accuracy(%)
Nearest Neighbor	92.85
Random Forest	88.57
Novel Min-Max classifier	98.57

By rendering the importance of emotion identification technique using Min-Max similarity classifier we attained an emotion identification efficiency of 98.57%.

D. Evaluation of facial features utilized for Emotion Identification

For the tests we utilized , a standard KDEF (Karolinska Directed Emotional Faces) dataset[10]. The dataset consists of face images of 70 individuals – out of which 35 male and 35 female. Each individuals face was captured expressing 6 various emotions– bliss, anguish, curiosity, despair, fury, hatred, and contempt Fig. 8. Also we captured every individuals face when the are said to be natural , having no emotions.

1. Detecting the region of the skin in the image

Fig. 8. The output of skin area identification utilizing the proposed algorithm

Images are treated in YCbCr space. To obtain the chrominance values (Cb and Cr) skin pixels considered and further verifying the proximity of a given pixel to the histogram center. It was a basic requirement to fix a threshold [11], which permits to identify pixels pertaining to facial skin. For this task, we need to obtain the histogram of human face colors, in the chromatic color space. We planned to utilize the EBM technique for skin identification.For the EBM technique, it9 is necessary to form a set of illustrations using skin region. Later 5% of the pixels, having the lowest average chrominance mixture, are discarded from the training set.

2. Detection of characteristic facial regions

The next phase in the technique of face detection is to utilize the Viola - Jones algorithm to determine facial regions, like eyes, eyebrows and lips[12].To enhance the efficacy of the technique, we planned to merge a previously defined technique for face, eyes, eyebrows and mouth identification shows in Fig. 9 using a classifier trained with the help of KDEF database.

3. Extraction of geometric features

In the initial phase the face regions are binarized using Otsu technique of optimal thresholding Fig. 10 [6] .We set a threshold value and segregate the pixels based on pixel values above and below the threshold. To enhance the result of binarization, we additionally use morphological operations .

Fig. 9. Detection of eyes, lips and eyebrows

Fig. 10. Outputs of binarization of sections of the face (top row - right eye, left eye, right eyebrow, bottom row - left eyebrow, mouth).

4. Extraction of anthropometric features

The Features from Accelerated Segment Test (FAST) algorithm was utilized to obtain the anthropometric features. The algorithm was selected, purely because of its good efficiency[4]. It can accurately detect the 4 points of the mouth, eyes (2 edges and the peak and the least points), and 2 points of the eyebrow . The overall output, of the entire task, is fetching sixteen points shows in Fig. 11 in the image, which later could be utilized to compute the anthropomorphic features.

Fig. 11. Detection of characteristic points on the image and example of extorted anthropometric features

5. *Results*

The classifiers were validated in Table II with additional faces, that were utilized in the training process[9-12].Three varied classifiers were verified: k-NN, SVM and Bagged Trees.

TABLE II

CLASSIFICATION EFFICIENCY FOR THE K-NN ,SVM, BAGGED TREE CLASSIFIER

Efficiency	K-NN	SVM	BAGGED TREE
Accuracy of categorization of six emotions	52.8%	55.9%	57.7%
Accuracy of the categorization of 2 emotions with the least performance(Bliss and curiosity)	92.7%	94.3%	95.9%
Accuracy of categorization of 2 emotions with the least efficiency(despair and disgust)	83.1%	83.5%	75.9%
Accuracy of categorization of the 2 emotions that frequently complicate among themselves(despair and curiosity)	78.7%	69.7%	75.3%

III. CONCLUSION

The paper that we analyzed utilizing interest points, depends on a correlation amidst interest points of face images having an emotion and of that having a neutral expression. The final gradients were categorized and envisioned in similarity plots that would be finally treated using a classifier. The eventual objective is to have a realistic facial emotion recognition system. The paper depending on Facial Emotion Recognition utilizing Min-Max Similarity Classifier is elementary, effortless and have obtained emotion recognition which surpasses the accomplishment of the previous techniques for the JAFFE database having an accuracy of 98.57%. The objective of the algorithm is to increase the emotion recognition efficiency by curbing the feature outliers and discarding the intensity offsets. Depending on the examination of facial features, a system that permits extraction of geometrical and anthropometric features was formed from the face images. Based on this few tasks were carried such as correlation of the working of varied classifiers that are used in the emotion recognition process. The nearest neighbours (k-NN), the support vector machine (SVM) and (Bagged Trees) decision trees classifiers, utilizing the bootstrap aggregation method were compared. For 6 various emotions (bliss, anguish, curiosity, despair, fury, hatred).we have achieved an average classification accuracy of 57.7% and an average classification accuracy of 95.9% for 2 emotions (bliss, curiosity).

IV. FUTURE ENHANCEMENT

Future work aims for an automated version of rectifying the flaws that are caused due to head movements. These head movement's leads to mistakes in computations and also affects the interest points. Our automated design should be in such a way that , even if the person raises his chin or lightly turns his head, interest points will never be disturbed. The memory used should be low. We can also introduce an automatic music player system based on facial emotions.

REFERENCES

[1] Marcin Kołodziej, Andrzej Majkowski, Remigiusz J. Rak, Paweł Tarnowski, Tomasz Pielaszkiewicz," *Analysis of facial features for the use of emotion recognition*",19th International Conference Computational problems of Electrical Engineering, 2018.

[2] Victor M. Álvarez, Ramiro Velázquez, Sebastián Gutiérrez, Josué Enriquez-Zarate," *A Method for Facial Emotion Recognition Based on Interest Points*", International Conference on Research in Intelligent and Computing in Engineering,2018.

[3] Olga Krestinskaya and Alex Pappachen James," *Facial Emotion Recognition using Min-Max Similarity Classifier*", International Conference on Advances in Computing, Communications and Informatics(ICACCI), 2017.

[4] M. Chetouani, J. Cohn, and A. A. Salah, Eds., *Human Behavior Understanding*: 7th International Workshop, HBU 2016, Amsterdam, The Netherlands, October 16, 2016, Proceedings. Springer International Publishing, 2016.

[5] B. Kim, J. Roh, S. Dong and S. Lee, "Hierarchical committee of deep convolutional neural networks for robust facial expression recognition", Journal on Multimodal User Interfaces, pp. 1–17, 2016.

[6] G. Levi and T. Hassner, "*Emotion recognition in the wild via convolutional neural networks and mapped binary patterns*", in Proc. of ACM International Conference on Multimodal Interaction, Seattle, WA, USA, pp. 503-510, 2015.

[7] H. Alshamsi, V. Kepuska and H. Meng, "*Real time automated facial expression recognition app development on smartphones*", in Proc. of IEEE Annual Information Technology, Electronics and Mobile Communication Conference, Vancouver, BC, Canada, pp. 384-389, 2017.

[8] X. Wang, X. Liu, L. Lu, and Z. Shen, "*A new facial expression recognition method based on geometric alignment and lbp features*," in 2014 IEEE 17th International Conference on Computational Science and Engineering, Dec 2014, pp. 1734–1737.

[9] S. Kamal, F. Sayeed, M. Rafeeq, and M. Zakir, "*Facial emotion recognition for human-machine interaction using hybrid dwt-sfet feature extraction technique*," in Cognitive Computing and Information Processing (CCIP), 2016 Second International Conference on. IEEE, 2016, pp.

[10] Elfenbein, H.A., Mandal, M.K., Ambady, N., Harizuka, S. and Kumar, S., 2002. Cross-cultural patterns in emotion recognition: Highlighting design and analytical techniques. Emotion, 2(1), p.75.

[11] Izard, C.E., 1994. Innate and universal facial expressions: evidence from developmental and cross-cultural research.

[12] Subash Kumar and Nagarajan V,"Local contourlet tetra pattern for image retrieval", Springer Journal of Signal Image and Video Processing, https://link.springer.com/article/10.1007/s11760-017-11971