Facial Emotion Recognition Based on Biorthogonal Wavelet Entropy, Support Vector Machine, and Stratified Cross Validation

YASH GUPTA, PRABHAT SINGH, AKASH KUMAR, SAURABH SINGH AND RAHUL KUMAR

Indian Institute of Information Technology,
Allahabad, 211015,
U.P, India.

ABSTRACT: Human emotions are not only important in relations but also play an important role in the way which we interact with the computer. Many algorithms have been developed for recognizing human emotions but they have not been up to the mark in terms of accuracy. In this project we aim to develop a Facial Emotion Recognition[1] model based on Biorthogonal Wavelet Entropy as a Feature Extractor and Support Vector Machines as a classifier, and test it on various datasets including benchmark datasets and a dataset created on our own. We trained and tested our software on JAFFE dataset and accuracy was found to better than earlier emotion recognition techniques.

INDEX TERMS: Facial emotion recognition, facial expression, biorthogonal wavelet entropy, support vector machines.

I. INTRODUCTION

Human emotions are complicated, and very difficult to envision. However, psychological studies have proven that emotions play a critical role in logical and intellectual behavior. Emotions are a complex mental state that drives a person to react positively or negatively to a stimulus.

Emotions help in different types of communications, like, verbal or non-verbal, peer to peer or community. But, emotions are very much short-lived as compared to other effects, namely behavior or mood. Nevertheless, emotions have been considered as signs of intelligence. By adding meaning to syntax, emotions make communication interactive and understandable. Even though it is very difficult for a machine to understand the human emotional state. However, since the onset of artificial intelligence[14],[15] there have been differences between the researchers defining this term. This is the reason that till now the machines are not as smart as human beings in terms of synergy.

Emotional recognition[2] is the process of identifying human emotions, largely facial expressions. Both of these things are done by the person himself but computational methods have also been developed. Recently, Human-Robot interaction have come into the limelight. Socializing robots with humans, comprehending facial expressions and visual hints of a human have been in great demand. It permits the robot to understand the manifestations of humans, which in turn can increase its effectiveness in fulfilling various tasks. This acts as a measurement system for behavioral science. Socially intelligent software tools can be achieved. Emotions can be reflected through voice, gestures, facial expression etc. A person's physical condition may also change, as he expresses different feelings. There are six major emotions i.e. happiness, sadness, fear, hatred, anger and wonder. Research in this area is still evolving, because so far in the developed systems, the feelings are not detected since factors and thus we have less accuracy.

One possible way to understand this complicated behavior is to understand human emotions. Human emotions can be reflected by many patterns such as facial expressions, gestures, muscular movements, speech tone, dermal activity etc. In spite of these important settings, psychological indicators were given more attention to the identification of human emotions among stimulating computing researchers.

II. SUBJECTS

We tested our software on the Japanese Female Facial Expression(JAFFE)[20] dataset. It contains 213 images of different emotions.
We also created our own dataset for testing.
We enrolled 10 subjects with ages from eighteen to twenty one of our institute. Each subject was requested to lay seven facial expressions (happy, sadness, surprise, anger, disgust, fear, and neutral). We photographed each facial expression of each subject ten times by Canon DSLR .

III. FEATURE EXTRACTION

A. BIORTHOGONAL WAVELET ENTROPY

Wavelet is a small portion of a wave, which is continuously decaying. It can be defined over a finite interval and having an average value of zero. Wavelet transform decomposes signal into a set of basic functions (wavelets). Wavelets are obtained from a prototype wavelet $\psi(t)$. Fourier transform is a powerful tool for data analysis because it do not represent abrupt changes efficiently which is required in the image analysis , Therefore to analyze abrupt changes we need a function (wavelet) which are well localized in time and space . So we use wavelet transform. Biorthogonal Wavelet Transform is generalized from Orthogonal Wavelet Transform[6] by using an extra set of low pass and high pass filters[5] for decomposition and reconstruction. Biorthogonal Wavelets allows more freedom[7] than orthogonal[4] and help in creating symmetric wavelet functions.

Figure 1: Samples of the dataset.

In the second correction, we mixed Shannon Entropy[8],[9] with BWT, and thus a proposal of multiscale[3] feature extraction technique called Biorthogonal Wavelet Entropy (BWE) was made. Shannon Entropy was demonstrated on each coefficient sub band, which is considered as one gray-level Image. The Shannon entropy is implemented with following equation:

$$H(X) = -\sum_{i=1}^{n} p(x_i) \log_b p(x_i)$$

Here X is a coefficient subband. x_i is the i-th gray-level of X. p represents the probability density function. n is the total number of gray-values. H is the entropy operation.

Figure 3 shows the block diagram shows the calculation of a 2-level BWE in the first stage, a facial expression image.

IV. CLASSIFICATION

A. SUPPORT VECTOR MACHINE

Support Vector Machines are machine learning algorithms which take a set of data (training data) as input and outputs the choicest hyperplane which classifies the given input data into two classes, moreover, the generated hyperplane is used to categorize new data. The support vectors are a data points of the training data that are closest to the optimal hyperplane, if these data points are removed then they will alter the position of the dividing hyperplane. Because of this, these support vectors are considered critical elements of the data set which help in building the hyperplane. While generating this hyperplane the margin between these support vectors is maximized, so that the new data has the maximum likelihood of falling in the correct class. The generated hyperplane is n-dimensional, where n is the number of features in the training vectors. Some use cases of SVMs are classification, regression, outlier detection and clustering. We use one of these functions namely, regression, which identifies which side of the hyperplane the new input data will occupy.

Support Vector Machines are a relatively new form of supervised machine learning, compared to neural networks. SVMs are great for small data sets that contain few outliers. Whereas, other machine learning algorithms such as deep neural networks, use large sets of data, but always give very robust models. The decision of which classifier to use depends upon the size of the dataset and the general complexity of the problem.

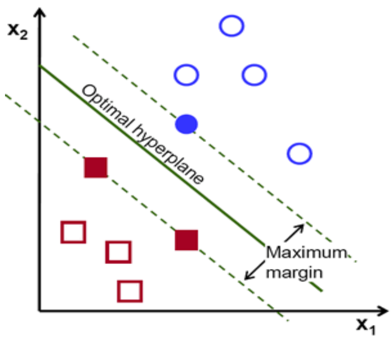

Figure 2: Optimal Hyperplane dividing the 2 classes [19]

In machine learning, we generally optimize an objective function, and in order to optimize an objective function we should minimize the error function. In SVMs, the use hinge loss as the error function.

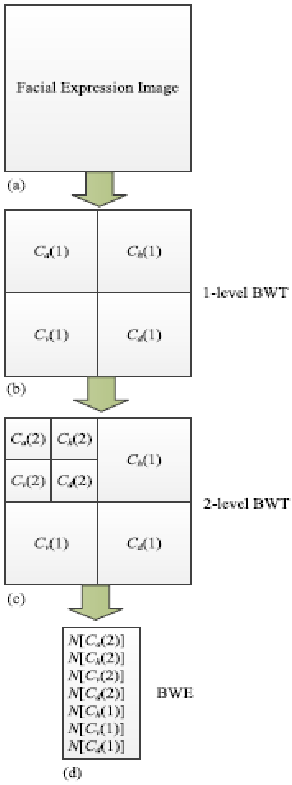

Figure 3: Block Diagram of a 2-level BWE: (a) A facial expression image; (b) 1-level BWT; (c) 2 level BWT; (d) Shannon entropy calculation.[18]

B. MULTICLASS. SUPPORT VECTOR MACHINE

Generally, SVMs are meant for 2-class problems.[10] But here we need to classify the emotions into 7 classes namely : Anger, Disgust, Fear, Happy, Neutral, Sad, Surprise. Thus, we split this problem into 7 different binary classifiers.[12] We apply the winner-takes-all(WTA)[11] approach here to find the best suited emotion for the given image. Figure 4 shows an example of a 7-class SVM with a score associated with every SVM, the one with the maximum score is chosen as the result.

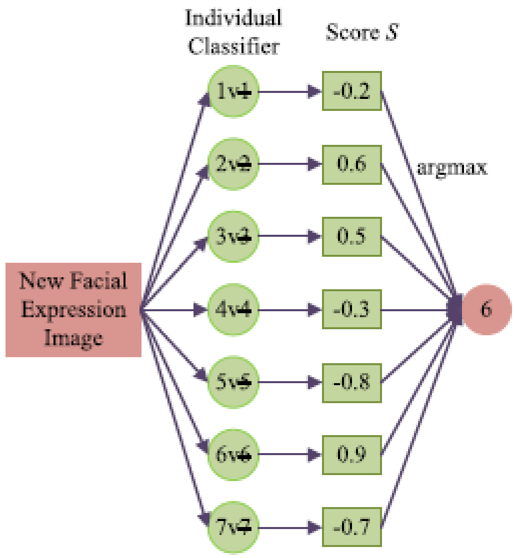

Figure 4: Diagram of a 7-class SVM using WTA technique.[18]

V. RESULTS AND DISCUSSIONS

We started with training our system on all seven emotions on JAFFE dataset. We have used ten fold stratified cross validation to validate our system. 213 image models contains 30 images for ANGER , 29 images for DISGUST , 32 images for FEAR , 31 images for HAPPY , 30 images for NEUTRAL , 31 images for SADNESS and 30 images for SURPRISE.

Using JAFFE dataset, we first trained our software on all seven emotions. Then there were similar results for disgust and anger so we trained on six emotions without disgust.

We get an Overall Accuracy of 79.3% after testing aour system on the JAFFE dataset.

Following are the results when we train our software on JAFFE dataset.

Table 1 : Each emotion accuracy on JAFFE dataset

Emotion	Accuracy
Anger	77.4%
Disgust	79.2%
Fear	70.3%
Happy	85.2%
Neutral	89.3%
Sad	88.9%
Surprise	69.8%
All 7 Emotions	**79.3%**

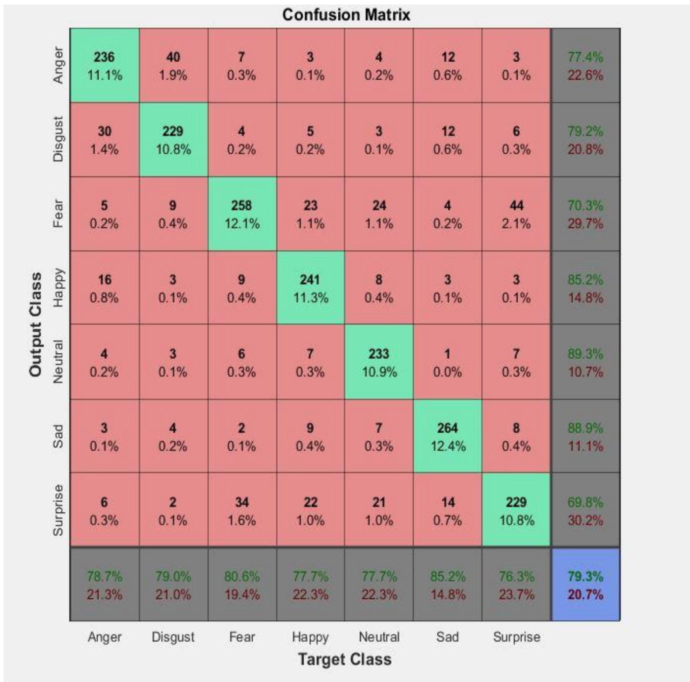

Figure 5: Confusion matrix for training on JAFFE dataset with 7 emotions

From the above confusion matrix we can observe that the results of the Anger and Disgust emotion are overlapping.

So, we found out the overall accuracy after removing the disgust emotion jumps to 81.7%. Interestingly, there is a significant rise in the accuracy of the anger emotion.

Table 2: Each emotion accuracy on JAFFE dataset (Without Disgust)

Emotion	Accuracy
Anger	86.9%
Fear	75.0%
Happy	89.3%
Neutral	81.1%
Sad	84.1%
Surprise	75.5%
All 6 Emotions	**81.7%**

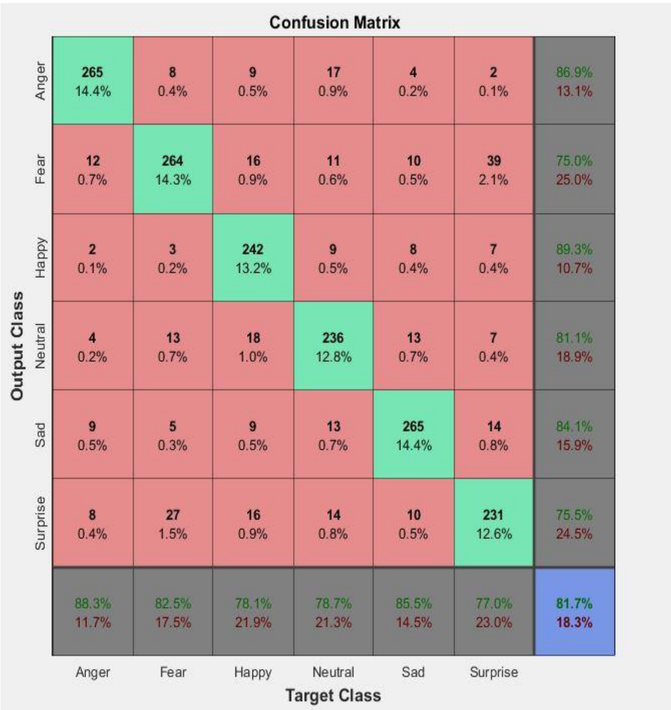

Figure 6: Confusion matrix for training on JAFFE dataset with 6 emotions

We created our own dataset by clicking images of our group members. Total number of images we used are 140 consisting of all seven emotions. Testing for our own dataset on our software lead to accuracy of 88.7%.

Most applications of emotion recognition examine static images of facial expressions. One such technique use Convolutional Neural Networks (CNN)[21] for emotion recognition in real time with a video input stream. It's accuracy was found to be 57.1% on JAFFE Dataset. Whereas the technique we used for classification is Support Vector Machine (SVM) which leads to accuracy of 79.3% on same JAFFE dataset.

VI. CONCLUSION

In this study, we proposed a new facial recognition system based on facial expression image. Our system selects biorthogonal wavelet entropy and multiclass support vector machine. Sometimes user emotion is not unique, it is mixture of two or more emotions. Fuzzy logic[13] can be implemented and the algorithm and result is Mixed Emotion. Our Project can be implemented in a music system[16],[17] which detect person's emotion and then play a suitable song accordingly. For example if a person is sad, it will play a happy song and if a person is happy, a party song is played. It is observed that physiological features like heart-beat rate can classify fear better than other features. Blood pressure is generally reduces in case of sadness. Hence inclusion of physiological features may lead to better accuracy.

REFERENCES

[1] S. Bernaerts, E. Berra, N. Wenderoth, and K. Alaerts, ``Influence of oxytocin on emotion recognition from body language: A randomized placebo-controlled trial," Psychoneuroendocrinology, vol. 72, pp. 182-189, Oct. 2016.

[2] M. G. Calvo and D. Beltrán, ``Brain lateralization of holistic versus analytic processing of emotional facial expressions," Neuroimage, vol. 92, pp. 237-247, May 2014.

[3] R. C. Guido, ``Practical and useful tips on discrete wavelet transforms [tips & tricks]," IEEE Signal Process. Mag., vol. 32, no. 3, pp. 162-166, May 2015.

[4] Y. C. Guo, L. L. Hu, and R. Ding, ``Orthogonal wavelet transform weighted multi-modulus blind equalization algorithm based on quantum particle swarm optimization," Acta Phys. Sin., vol. 61, Mar. 2012, Art. no. 054304.

[5] X.-G. Zhu, B.-B. Li, and D.-F. Li, ``Orthogonal wavelet transform of signal based on complex B-spline bases," Int. J. Wavelets Multiresolution Inf. Process., vol. 10, no. 6, Nov. 2012, Art. no. 1250054.

[6] H. Sun, J. Wang, and R. Shao, ``Depressing defocusing effect in microscopy by bi-orthogonal wavelet transform," Particuology, vol. 10, no. 5, pp. 644-647, Oct. 2012.

[7] O. Prakash and A. Khare, ``Tracking of moving object using energy of biorthogonal wavelet transform," Chiang Mai J. Sci., vol. 42, no. 3, pp. 783-795, Jul. 2015.

[8] K. H. Norwich, ``BoltzmannShannon entropy and the double-slit experiment," Phys. A, Statist. Mech. Appl., vol. 462, pp. 141-149, Nov. 2016.

[9] W. S. Nascimento and F. V. Prudente, ``Study of Shannon entropy in the context of quantum mechanics: An application to free and contend harmonic oscillator," Quimica Nova, vol. 39, no. 6, pp. 757-764, Jul. 2016.

[10] T. Genewein and D. A. Braun, ``Bio-inspired feedback-circuit implementation of discrete, free energy optimizing, winner-take-all computations," Biological, vol. 110, no. 2, pp. 135-150, Jun. 2016.

[11] Y. Zhang and L. Wu, ``Classification of fruits using computer vision and a multiclass support vector machine," Sensors, vol. 12, no. 9, pp. 12489-12505, 2012.

[12] M. Kim and C. M. Twigg, ``Rank determination by winner-take-all circuit for rank modulation memory," IEEE Trans. Circuits Syst. II, Express Briefs, vol. 63, no. 4, pp. 326-330, Apr. 2016.

[13] Y. Chen et al., ``Pathological brain detection by wavelet-energy and fuzzy support vector machine," in Proc. 8th Int. Symp. Comput. Intell. Design (ISCID), Hangzhou, China, Dec. 2015, pp. 409-412.

[14] S. Balochian, ``Artificial intelligence and its applications," in Mathematical Problems in Engineering. USA: Hindawi Publishing Corporation 2014, Art. no. 840491.

[15] P. Agarwal, ``Artificial intelligence and its applications 2014," in Mathematical Problems in Engineering. USA: Hindawi Publishing Corporation 2016, Art. no. 3871575, doi: 10.1155/2016/3871575.

[16] Y. Zhang, ``GroRec: A group-centric intelligent recommender system integrating social, mobile and big data technologies," IEEE Trans. Serv. Comput., vol. 9, no. 5, pp. 786795, Sep./Oct. 2016.

[17] Y. Zhang, M. Chen, D. Huang, D. Wu, and Y. Li, ``Doctor: Personalized and professionalized medical recommendations based on hybrid matrix factorization," Future Generat. Comput. Syst., vol. 66, pp. 3035, Jan. 2017.

[18] Y. Zhang, Z. Yang, H. Lu, X. Zhou, P. Phillips, Q. Liu and S. Wang, "Facial Emotion Recognition Based on Biorthogonal Wavelet Entropy, Fuzzy Support Vector Machine, and Stratified Cross Validation", IEEE Access, vol. 4, 2016.

[19] "Introduction to Support Vector Machines — OpenCV 2.4.13.3 documentation", Docs.opencv.org, 2017.[Online]

[20] "Facial Expression Database: Japanese Female Facial Expression (JAFFE) Database", Kasrl.org, 2017. [Online].

[21] D. Duncan, G. Shine and C. English, "Facial Emotion Recognition in Real Time", 2017.